ANÁLISE ACÚSTICA DA CONSOANTE /P/ POR FALANTES BILÍNGUES DE INGLÊS ESTADUNIDENSE E PORTUGUÊS BRASILEIRO

OTAVIO AUGUSTO RODRIGUES BERNARDO SILVA

Editora Poisson

Otavio Augusto Rodrigues Bernardo Silva

Análise acústica da consoante /p/ por falantes bilíngues de inglês estadunidense e português brasileiro

1ª Edição

Belo Horizonte

Editora Poisson

2024

Editor Chefe: Dr. Darly Fernando Andrade

Conselho Editorial

Dr. Antônio Artur de Souza – Universidade Federal de Minas Gerais
MSc. Davilson Eduardo Andrade
Dra. Elizângela de Jesus Oliveira – Universidade Federal do Amazonas
MSc. Fabiane dos Santos
Dr. José Eduardo Ferreira Lopes – Universidade Federal de Uberlândia
Dr. Otaviano Francisco Neves – Pontifícia Universidade Católica de Minas Gerais
Dr. Luiz Cláudio de Lima – Universidade FUMEC
Dr. Nelson Ferreira Filho – Faculdades Kennedy
MSc. Valdiney Alves de Oliveira – Universidade Federal de Uberlândia

International Cataloging Data in Publication (CIP)

B523 Silva, Otavio Augusto Rodrigues Bernardo

Análise acústica da consoante /p/ por falantes bilíngues de inglês estadunidense e português brasileiro; Belo Horizonte - MG: Poisson,2024

Formato: PDF
ISBN: 978-65-5866-408-6
DOI: 10.36229/978-65-5866-408-6
Modo de acesso: World Wide Web
Inclui bibliografia

1. Análise acústica 2. Falantes bilíngues
I. Bernardo, Otavio Augusto Rodrigues

CDD-414

Sônia Márcia Soares de Moura – CRB 6/1896

O conteúdo deste livro está licenciado sob a Licença de Atribuição Creative Commons 4.0.
Com ela é permitido compartilhar o livro, devendo ser dado o devido crédito, não podendo ser utilizado para fins comerciais e nem ser alterada.

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos seus respectivos autores.

Esse e outros títulos podem ser baixados gratuitamente em www.poisson.com.br
Entre em contato pelo contato@poisson.com.br

Autor

Otavio Augusto Rodrigues Bernardo Silva

Doutorando em Linguística na Universidade Federal do Paraná (UFPR). Cursou ensino médio técnico integrado em Eletrônica e licenciatura em Letras Inglês na Universidade Tecnológica Federal do Paraná (UTFPR). Cursou mestrado em Estudos Linguísticos na Universidade Federal do Paraná (UFPR). Possui experiência na área de ensino de inglês como língua estrangeira, transcrição e tradução. Desenvolveu projetos de docência interdisciplinar através do PIBID em colégios da rede estadual de ensino, é professor particular de língua estrangeira, professor em institutos de idiomas e criador de materiais didáticos e cursos online de língua inglesa.

Prefácio

Este trabalho foi desenvolvido em tempos de pandemia, distanciamento social, inseguranças e tragédias. Desde o princípio, nada saiu como havia sido idealizado: desde as disciplinas de mestrado feitas remotamente, com equipamentos e internet deficitários, até aceitar bancos de dados gentilmente cedidos para realizar as análises, em vez de coletar o meu próprio. Mesmo assim, o trabalho é uma sólida análise em diversas características acústicas da consoante /p/ produzida por falantes de duas nacionalidades, brasileira e estadunidense, e serve como uma válida contribuição para pesquisadores provenientes não apenas das áreas de fonética e fonologia, mas também estudiosos interdisciplinares de áreas como a (linguística) forense e das teorias da complexidade.

SUMÁRIO

1. INTRODUÇÃO 08

2. TEORIA DOS SISTEMAS DINÂMICOS COMPLEXOS 11

2.1. COMPLEXIDADE E LÍNGUA 11

2.2. PERCEPÇÃO E PRODUÇÃO DOS SONS DE L2 16

3. FONÉTICA FORENSE 19

3.1. FONÉTICA E FONÉTICA ACÚSTICA 19

3.2. FONÉTICA FORENSE E COMPARAÇÃO DE LOCUTORES 20

4. CONSOANTES OCLUSIVAS 24

4.1. PARÂMETROS ARTICULATÓRIOS E ACÚSTICOS DE OCLUSIVAS 24

4.2. DEFINIÇÃO DE OCLUSÃO E VOT 26

4.2.1. OCLUSÃO 26

4.2.2. VOT 29

4.3. REVISÃO DE TRABALHOS: BILINGUISMO E VOT 33

5. METODOLOGIA 40

5.1. EXPERIMENTO EXPLORATÓRIO 41

5.2. O ESTUDO 43

5.3. HIPÓTESES 44

6. RESULTADOS 46

6.1. RESULTADOS QUANTITATIVOS 46

6.1.1. MÉDIAS DE DURAÇÃO DE OCLUSÃO 48

6.1.2. MÉDIAS DE DURAÇÃO DE VOT 54

6.1.3. MÉDIAS DE DURAÇÃO TOTAL DE CONSOANTE 63

6.1.4. COMPARAÇÃO DE TAREFAS: TEXTOS E SENTENÇAS-VEÍCULO 68

6.2. ANÁLISE QUALITATIVA 69

6.3. VERIFICAÇÃO DAS HIPÓTESES 73

7. CONSIDERAÇÕES FINAIS 75

REFERÊNCIAS 78

APÊNDICES 81

Resumo: A presente pesquisa investiga a produção de sete estadunidenses e sete brasileiras bilíngues falantes de inglês estadunidense e português brasileiro. A análise busca cobrir a duração do período de oclusão, VOT e a duração total da consoante (junção da oclusão e do VOT, medido com o burst) de /p/ em posição silábica inicial. A visão de língua adotada é a de Sistema Dinâmico Complexo (SDC), que consiste em uma teoria interdisciplinar proveniente da matemática e que é adotada por diversos campos de estudo como a Física e a Química (BARANGER, 2000). Na Linguística, a complexidade pode ser utilizada como metáfora para o desenvolvimento linguístico: a língua possui agentes interdependentes, com múltiplas escalas de interação, emergência de padrões e comportamentos e relações de caos e não caos. Além disso, trabalhos como os de Lopes e Baumgartner (2019) descrevem a língua inglesa como a mais falada no mundo e a língua que cada vez mais brasileiros se interessam em desenvolver, justificando pesquisas com bilinguismo. Busca-se também contribuir para a área de Fonética Forense por meio do ramo de comparação de locutores. É importante identificar aspectos únicos da fala de uma pessoa e o estudo dos períodos como oclusão e VOT de uma consoante oclusiva podem ajudar na tarefa. Como resultados entre as nacionalidades foram encontradas diferenças significativas entre brasileiras e estadunidenses nas produções de durações absolutas e relativas de VOT em inglês, durações absolutas e relativas da oclusão em inglês, duração relativa da oclusão em português e duração relativa da consoante em português. Os resultados dentro de cada grupo de nacionalidade apontaram que estadunidenses produzem durações de VOTs e durações absolutas de oclusões, além de durações de VOTs e durações relativas de consoantes de forma diferente em português e inglês, já brasileiras produzem durações absolutas e relativas de VOTs, durações relativas de oclusões e durações absolutas de consoantes diferentemente nas duas línguas. Uma análise qualitativa que comparou espectrogramas também foi desenvolvida. Aspectos como a diferença significativa nas durações entre ambas as nacionalidades em inglês e o maior desvio-padrão nas durações de cada nacionalidade em sua respectiva língua estrangeira podem colaborar para o ramo da comparação de locutores.

Palavras-chave: Sistemas Dinâmicos Complexos; Fonética Forense; Fonética Acústica; Consoantes Oclusivas; Voice Onset Time; .

CAPÍTULO

01 INTRODUÇÃO

Gostaria, primeiramente, de fazer um breve relato das experiências acadêmicas até aqui, para, então, discorrer um pouco sobre a experiência do mestrado e apresentar a divisão do texto. Fui instigado a participar de pesquisas desde o começo da graduação, primeiramente por meio da disciplina de Introdução à pesquisa em Linguagem. Durante quatro semestres entre os anos de 2016 e 2019, participei (formal ou informalmente) de seis grupos de pesquisa ofertados pelo curso recém-inaugurado de Licenciatura em Letras Inglês da Universidade Tecnológica Federal do Paraná (UTFPR). A partir desses grupos e de projetos externos, tive a oportunidade de desenvolver pesquisas iniciais nas áreas de fonética acústica (vogais paragógicas) e gêneros textuais (modelo didático do gênero Histórias em Quadrinhos), apresentando os trabalhos em formato de pôster em 2017, no Congresso da Abralin. Por intermédio de minha orientadora de TCC, Maria Lúcia de Castro Gomes, fui incentivado a continuar a pesquisar. Apresentei mais dois trabalhos nos anos de 2018 e 2019 (em Florianópolis e Maceió, respectivamente), além de publicar um artigo em conjunto (Revista *Language and Law*[1]) e receber uma bolsa para desenvolver o TCC. Ademais, tive a oportunidade de ser estagiário do Departamento de Línguas Estrangeiras Modernas por quase dois anos, bem como participar por um ano do Programa Institucional de Bolsa de Iniciação à Docência (PIBID) interdisciplinar, no qual publicamos dois trabalhos com colegas da Química, Matemática, Português e Física, sem contar a experiência profissional docente e as experiências dos estágios. Tudo isso foi essencial para conseguir chegar ao mestrado, que rendeu o presente texto, desenvolvido nos anos de 2020 e 2021. Com relação ao desenvolvimento da dissertação, afirmo que foi muito difícil produzir durante a pandemia. A situação do Brasil só piorou desde março de 2020, o que dificultou mais o que já estava desafiador. A ideia inicial do trabalho era estudar consoantes oclusivas e VOT, mas o estudo final foi de apenas uma consoante (no caso /p/) em vários outros parâmetros (durações de oclusão e VOT, além do total da consoante). O desejo por estudar oclusivas vem desde a primeira vez que estudei VOT na universidade, no terceiro período, e imaginei que uma pesquisa de bilinguismo[2], enxergando a língua como SDC poderia contribuir para todas as áreas investigadas. Acredito que os dados de durações de /p/ por estadunidenses bilíngues residentes no Brasil sejam a principal contribuição do trabalho. Passo a falar, especificamente, da pesquisa nos parágrafos a seguir.

De acordo com o portal estatístico *Statista* (2021), mais de um bilhão e 300 milhões de pessoas no mundo falam inglês em 2021. Esse número conta com falantes nativos e falantes de inglês como segunda língua. Isso faz o inglês ser o idioma mais popular do mundo, seguido do Mandarim, com um bilhão e 100 milhões de falantes e o Hindi, com 600 milhões de falantes. Lopes e Baumgartner (2019) trazem dados mundiais que se

[1]Nas referências: Bernardo-Silva e Gomes (2019).
[2]Na atual pesquisa, o bilinguismo é entendido como a coexistência de dois subsistemas linguísticos em um falante. Pressupõe-se que exista interações multidirecionais entre os subsistemas.

referem a 1,5 bilhão de falantes de inglês em 2002, sendo somente 350 milhões destes falantes nativos. Conclui-se, portanto, que o número de falantes de inglês como língua estrangeira (por Lopes e Baumgartner entendida como língua aprendida fora de países em que o inglês é a primeira língua) ou segunda língua (pelos autores entendida como a língua de falantes que aprendem mais de uma língua em seu próprio país, como na Índia) é muito maior do que o número de falantes de inglês como língua materna. Visões de língua internacionalizadoras como Inglês como Língua Franca (ELF) são desenvolvidas a partir desse panorama (LOPES; BAUMGARTNER, 2019).

Segundo levantamento de dados feito pelo *British Council* (2014), apenas 5% da população brasileira indicava conhecimento de inglês. Em geral, o aprendizado do idioma sempre esteve disponível apenas para as classes abastadas. O estudo indicava que as classes média e média alta se interessavam cada vez mais pelo idioma, principalmente pela possibilidade de melhor empregabilidade. Tal procura crescente pelo estudo de inglês como L2 ou segunda língua[3] justifica pesquisas que consideram a importância da área de bilinguismo.

Adquirir conhecimento sobre os processos de desenvolvimento de uma segunda língua pode ser importante para diversas áreas do conhecimento, como, por exemplo, a Psicologia (especificamente psicolinguística) e a Neurologia (neurolinguística). Neste trabalho, buscar-se-á também colaborar para o ramo da fonética forense, especialmente para a comparação de locutores, que consiste em buscar características que individualizam falantes por meio da análise auditiva e acústica. Um exemplo de aplicação dos conhecimentos da área da fonética forense é observado na comparação de uma gravação de voz (que pode ser evidência de um crime) com a fala de um suspeito, através de recursos da fonética acústica e auditiva. A área de estudo da fonética forense pode se interseccionar com o estudo da língua como Sistema Dinâmico Complexo (SDC), uma vez que a complexidade se encontra além de visões simplificadoras de língua/fala. Os órgãos responsáveis pela fala, as emoções e eventuais disfarces de locutores são agentes de um sistema complexo e interconectado. A grande variabilidade de produções analisadas durante a pesquisa indica que cada informante também produz sons de forma variável e não linear. A visão de língua dentro do paradigma da Teoria dos Sistemas Dinâmicos Complexos (TSDC)[4] foi adotada nesta pesquisa, portanto.

Considerando o inglês como a língua mais falada do mundo por conta de seu *status* de língua global/franca/internacional entre falantes nativos e não nativos, incluindo brasileiros, a visão de língua como SDC e implicações para a fonética forense, apresento o objetivo principal do trabalho: analisar a duração de oclusão, VOT e duração total da consoante na produção de /p/ por falantes brasileiras e estadunidenses em português e inglês. Os objetivos específicos são:

- verificar, em uma análise intergrupos, se as durações de Oclusão, VOT e duração total da consoante são diferentes em cada uma das línguas;
- verificar, em uma análise intragrupo de estadunidenses, se as durações de Oclusão, VOT e duração total da consoante são diferentes em cada uma das línguas;

[3]No presente trabalho, os termos L2, segunda língua e língua estrangeira são usados como sinônimos de uma língua que não é a materna, evitando embarcar nas discussões teóricas (alguns conceitos já foram apresentados) sobre as terminologias.
[4]Termo não concorrente com a definição da teoria, que é Teoria dos Sistemas Dinâmicos Complexos (TSDC). As discussões sobre a terminologia se encontram no Capítulo 2.

- verificar, em uma análise intragrupos de brasileiras, se as durações de Oclusão, VOT e duração total da consoante são diferentes em cada uma das línguas;
- verificar, através da comparação de durações de oclusão, VOT e duração total da consoante produzidas pelos dois grupos nas duas línguas, se existem diferenças significativas de acordo com a coleta de dados efetuada, ou seja, por leitura de sentenças-veículo ou leitura de texto;
- discutir quais são as implicações desta pesquisa para a fonética forense.

A pesquisa analisa a fala de 14 informantes, sete brasileiras e sete estadunidenses. Dois instrumentos diferentes de coleta de dados foram utilizados: um texto nas duas línguas (lidos uma vez em cada língua) e sentenças-veículo nas duas línguas (lidas três vezes em cada língua). Os parâmetros quantitativos analisados foram a duração absoluta (em milissegundos) e relativa (tamanho em relação à porcentagem da palavra) do período de oclusão e do VOT de /p/, assim como a duração total da consoante (soma da oclusão e do VOT da consoante até o início da vibração das pregas vocais da vogal seguinte) em cada ocorrência nos textos e sentenças-veículo. Cabe adicionar que os valores do VOT da plosiva em alguns casos (principalmente em algumas repetições de brasileiras falando português) se aproximavam dos 10ms e ficavam muito difíceis de se dissociar do *burst* da oclusão. Por essa razão, o valor da duração do *burst* foi contabilizado junto a todos os valores de VOT analisados. Análises qualitativas foram feitas a partir de exemplos de espectrogramas de ambas as nacionalidades.

O texto está organizado da seguinte maneira: no Capítulo 2, apresento a visão de língua como SDC, oferecendo características inerentes aos sistemas e porque é possível relacionar a língua com essas características; no Capítulo 3, indico o que é a fonética (acústica e forense) e porque pesquisas nesse ramo são importantes; no Capítulo 4, ofereço uma descrição teórica do que são as consoantes oclusivas e dos trabalhos importantes que oferecem subsídios às análises de oclusiva, com foco nos momentos de oclusão e VOT; no Capítulo 5, detalho como foi elaborada a metodologia de pesquisa e qual foi o estudo exploratório realizado em 2020. Apresento também as hipóteses do trabalho; no Capítulo 6, descrevo como foi o estudo principal, os resultados obtidos, quantitativamente e qualitativamente, e, por fim, retomo e verifico as hipóteses; o Capítulo 7 é reservado às considerações finais.

CAPÍTULO

02 TEORIA DOS SISTEMAS DINÂMICOS COMPLEXOS

Este Capítulo se propõe a estudar a Teoria dos Sistemas Dinâmicos Complexos (TSDC) em suas origens e desdobramentos na Seção 2.1 e a mapear, em seguida, modelos de desenvolvimento linguístico como o *Speech Learning Model* (SLM) que é, frequentemente, utilizado por parte considerável da bibliografia do presente trabalho.

2.1. COMPLEXIDADE E LÍNGUA

Nesta seção, apresento, brevemente, o que é a Complexidade. Parto dos pressupostos teóricos de Sistemas Dinâmicos Complexos (SDC) e Sistemas Adaptativos Complexos (SAC) e justifico o porquê de adotar esses modelos para analisar dados acústicos.

A definição precisa da complexidade, de acordo com Baranger (2000) e Holland (2006), até a data de seus respectivos estudos, não estava bem consolidada. Baranger (2000) pontua que, quanto mais pesquisadores adentrarem na área, maiores as chances de uma definição ser encontrada. Isso explica o porquê da necessidade de tantas características basilares provenientes de tantos autores diferentes que se encontram na sequência.

Segundo Larsen-Freeman (2013), a teoria da complexidade tem por fonte as ciências físicas. A autora percebeu que os SDCs poderiam contribuir também para a Linguística Aplicada. Uma das dificuldades dos SDCs ao analisar dados é a dificuldade de delimitar as fronteiras entre fenômenos linguísticos, uma vez que a complexidade não se afina com a perspectiva de aprendizado linear (LARSEN-FREEMAN, 2013).

Holland (2006) afirma que é essencial definir como os agentes de um SDC agem. Os agentes de SDC atuam em paralelo, enviando e recebendo dados. Isso pode ser facilmente demonstrado com a língua: os falantes de uma língua, constantemente, enviam e recebem dados de fala. Além disso, o autor também adiciona que muitos sinais acontecem simultaneamente, o que também pode ser compreendido ao tomarmos uma língua como exemplo (em uma sala de aula, alunos podem estar praticando simultaneamente, e mesmo assim enunciados únicos serão gerados). Além disso, Holland traz que os agentes de um SDC operam em um sistema condicional (SE/ENTÃO), como em SE alguém falar *bom dia*, ENTÃO um interlocutor *pode* responder *bom dia.* Adaptações e evoluções também podem ser observadas em um SDC, como um aprendiz de segunda língua que aprende e incorpora novas estruturas linguísticas.

Segundo Baranger (2000), o conceito de caos é matemático. Este professor de Física explica, em termos apropriados para quem não é de sua área, a resistência à aderência dos conceitos da complexidade. Na Física, por exemplo, o paradigma do cálculo, instituído por Leibniz e Newton na era moderna, ocupava muito tempo na grade dos alunos e oferecia certas complicações que tiravam os alunos de zonas de conforto. Um

exemplo dessas complicações seria aceitar a característica imprevisível de um SDC. De acordo com Baranger (2000), o paradigma previsível do cálculo se tornou tão intricado que passou a ser adotado como verdade absoluta e os teóricos simplesmente não enxergavam situações em que, talvez, os conceitos não fossem soberanos em sua função descritiva. Como exemplo concreto de complexidade, Baranger (2000) traz os fractais, conceitualmente, definidos como figuras geométricas, cujos traços se repetem, indefinidamente, ao se analisar cada vez mais atentamente. Elementos naturais, como as ondas do mar e o corpo humano, também seriam fractais. Podemos olhar para a linguagem dessa mesma maneira: um verdadeiro universo de léxico, significados, significantes, fonemas, implicaturas, sintagmas e outros. Baranger (2000) traz seis propriedades de um SDC, em uma tentativa de enumerar suas características:

1) SDCs contêm muitos constituintes que interagem de forma não linear (estes constituintes aparentam ser os mesmos agentes trazidos por Holland);
2) Os agentes de um SDC são interdependentes;
3) A estrutura de um SDC abrange várias escalas;
4) Um SDC consegue fazer emergir comportamentos;
5) A complexidade envolve uma relação do caos com o não caos;
6) A complexidade envolve uma relação entre o cooperativo e o competitivo.

Aqui, proponho colocarmos os exemplos do autor para complexidade, e, em seguida, meus exemplos para ligarmos a Linguística com a complexidade.

Em (1), tem-se o argumento de que caos e complexidade não são sinônimos (há de se ter certas condições não caóticas para um sistema ser complexo, de outro jeito seriam apenas puramente caóticos). O conjunto geral de um SDC, portanto, é não linear. Na área de ensino de inglês como L2, isso fica fácil de se observar com alunos em uma mesma sala de aula: alguns entendem *Simple Past* muito mais rápido que outros; outros aprendem *Past Continuous* de forma mais lenta e em outro estágio de aprendizado.

Com relação ao segundo tópico, o autor exemplifica com a amputação de uma perna no corpo humano: as consequências de remover 10% de um SDC, certamente, afetariam outras estruturas (podemos pensar na mobilização do sistema imunológico para evitar infecções e no trabalho das células anticoagulantes para que não haja morte por hemorragia). O exemplo do corpo humano pode parecer um pouco extremo, então na Linguística podemos observar algo mais leve: como uma pessoa pode ter seu vocabulário alterado por viver em uma região diferente do país e influenciar outros falantes com seu vocabulário natal. Como evidência anedótica, destaco que passei a utilizar *tu* pela presença de gaúchos na sala de aula, bem como *mano* pela presença de paulistas.

No tópico (3), o autor enumera as escalas do corpo humano (órgãos, tecidos, células, por exemplo). Na Linguística, temos a língua, dialetos, fala, idioleto, e assim por diante.

Em (4), dá-se que, com o intuito de entender estados emergentes, necessita-se olhar para diversas escalas de um sistema. É impossível entender o fenômeno de andar ao olhar somente uma perna (são usados os quadris, tornozelos, pés, dedos, braços e outros). Na língua, gírias podem emergir de qualquer nicho e de repente serem repetidas por parcelas inteiras de uma população.

No quinto tópico, o autor enfatiza a existência de certo parâmetro de controle que evita o transbordamento do sistema para o puro caos (a chamada "borda/fronteira do caos", que dividiria o complexo do caótico). Do ponto de vista da língua, se cada um atribuísse o significante que quisesse a qualquer objeto, a comunicação seria impossível. Há, porém, várias divergências em dialetos diferentes da mesma língua (autocarro e ônibus, por exemplo).

No tópico (6), famílias do século XIX são tomadas como exemplo: elas precisavam cooperar em seu núcleo para prosperar e seu embate com outras famílias geravam futuros casamentos mais prósperos. Na língua, podemos perceber como vocábulos concorrentes disputam a preferência dos falantes.

Chan (2001), em um seminário voltado para sistemas de engenharia, também afirmou o caráter interdisciplinar dos SDCs. Pesquisadores da área investigam os processos pelos quais um SDC responde a *inputs* diversos e a alterações internas e externas. A autora afirmou também que os SDC se distinguem por uso extenso de computadores para fins de simulação e como ferramentas de pesquisa. Chan definiu a complexidade como o produto de inter-relações, interações e interconectividade entre elementos de um sistema e o ambiente. Novamente, é possível perceber a linguagem dentro de tal definição: falantes interagem entre si, em diferentes tipos de relações. De acordo com Chan, as principais características dos SDCs são:

a) Controle distribuído, ou seja, sem um mecanismo de controle centralizado que dita as regras do sistema;

b) Conectividade, que é explicada pelas inúmeras formas em que uma parte do sistema pode gerar alterações para as demais pelos elementos estarem interligados e serem afetados uns pelos outros;

c) Coevolução, que faz os elementos evoluírem baseados em relações prévias com agentes terceiros;

d) Dependência sensível às condições iniciais, que resulta no fato de que pequenas alterações na entrada (*input*) do sistema podem gerar grandes mudanças futuras no todo;

e) Emergência, ou o potencial de certo comportamento emergir de um sistema.

f) Distância do equilíbrio ou a capacidade de um SDC balancear ordem e caos para se manter vivo;

g) Estado de paradoxo, para resumir todas as relações divergentes de ordem/caos, cooperação/competição e ordem e desordem presentes em um SDC.

Rzevski (2012), em um artigo que discute o paradoxo da auto-organização / controle de um SDC, apresenta sete critérios para definir uma ocorrência de sistema. O autor critica, duramente, o paradigma newtoniano e argumenta que os SACs/SDCs são muito mais apropriados para se lidar com sistemas envolvendo seres humanos. Descrevo as definições de SAC de Rzevski (2012):

1) Interdependência, que afirma que um sistema é constituído de agentes interdependentes e que influenciam uns aos outros;

2) Autonomia, em que os agentes não são controlados por um mecanismo central;

3) Emergência, que define que o comportamento global do sistema depende da interação dos agentes;

4) Não equilíbrio, uma vez que o autor afirma que a alta ocorrência de eventos disruptivos não permite que um SDC entre em equilíbrio;

5) Não linearidade, que prevê que *inputs* pequenos podem gerar grandes perturbações no sistema;

6) Auto-organização, ou seja, a capacidade do sistema voltar a um estado equilibrado depois de um evento disruptivo;

7) Coevolução; um sistema sempre evolui com seu ambiente.

Após identificar as características apontadas por autor, proponho um cruzamento de informações para perceber em quais pontos Baranger, Chan e Rzevski se interseccionam, como é possível observar no Quadro 1:

QUADRO 1 – COMPARAÇÃO DE CONCEITOS ENTRE TEÓRICOS

Baranger	Chan	Rzevski	CONCEITO
(1)	(d)	(V)	Não linearidade
(2)	(b)	(I)	Interdependência
(4)	(e)	(III)	Emergência
(5) e (6)	(g)		Paradoxo
	(a)	(II)	Ausência de Controle Central
	(f)	(IV)	Não equilíbrio
	(c)	(VII)	Coevolução

FONTE: O autor (2021).

Os seguintes conceitos não possuem paralelos diretos com a listagem dos outros autores, porém não deixam de seguir a mesma lógica da complexidade:

3) Múltiplas Escalas, de Baranger, relacionado com os exemplos de fractais.

VI. Auto-organização, de Rzevski, relacionado com a capacidade do SDC voltar ao equilíbrio.

Certos linguistas estadunidenses representados no trabalho Larsen-Freeman *et al.* (2013) desenvolveram suas pesquisas em Sistemas Adaptativos Complexos ao longo dos anos – linguistas europeus, com destaque para os Países Baixos, também estudaram como o desenvolvimento da complexidade em seu continente com a denominação Sistemas Dinâmicos Complexos (SDC). Os autores, dos quais descreverei alguns trabalhos como de De Bot *et al.* (2007), Verspoor *et al.* (2008) e Lowie e Verspoor (2019), partiram das raízes provenientes da Matemática e adentraram as características de SDCs para aplicações na Linguística. De Bot *et al.* (2007) descreveram os SDCs como dependentes de condições iniciais, contendo interações entre agentes, subsistemas interconectados, emergência de estados atratores no desenvolvimento por meio do tempo e variações entre grupos e indivíduos. De Bot (2017) afirma que a conciliação da Linguística (Aplicada) com os SDCs começou a ser traçada em 1997 com a publicação de Larsen-Freeman, porém demorou um tempo para que houvesse atenção e a adesão de linguistas. O autor afirma que a mudança de paradigma e a alteração do foco de vertentes teóricas são processos inerentes às mudanças do campo e da comunidade científica. Em De Bot (2017), faz-se claro que a união entre pesquisadores de Sistemas Adaptativos Complexos e Sistemas Dinâmicos Complexos era importante, uma vez que as correntes hegemônicas na linguística eram de ordem gerativista e que colidia com a concepção de língua como SDC. Segundo a pesquisadora brasileira Albuquerque (2019), a postura mais conciliadora por parte dos

pesquisadores de Sistemas Adaptativos e Sistemas Dinâmicos gerou uma nova sigla, a TSDC (Teoria dos Sistemas Dinâmicos Complexos). Os TSDC abarcam as características de desenvolvimento no tempo dos aprendizes de L2, presença de competitividade e cooperação no aprendizado e variabilidade no desenvolvimento da língua, por exemplo. Castro *et al.* (2019) é um trabalho recente desenvolvido por pesquisadoras brasileiras e já se desenvolve sob uma ótica de TSDC, estudando dados de um aprendiz de português brasileiro natural do Chile. Partindo dessa perspectiva, as autoras incorporam características da complexidade como a noção que processos de *influência* entre as línguas variam ao longo do tempo.

Lowie e Verspoor (2019) discutem os problemas de ergodicidade nos estudos de desenvolvimento de língua estrangeira. De forma resumida, os autores explicam que um sistema ergódico pode ser representado por variações uniformes – que podem ser individuais ou coletivas – ao longo do tempo. O problema é que, de acordo com os autores, os mais de 50 anos de evolução no estudo da área de diferenças individuais no aprendizado foram conturbados. Nessa pesquisa, na qual os dados apresentados são de 22 participantes holandeses aprendizes de inglês, a questão da ergodicidade é *complexa* e este (ergodicidade) pode não ser o caso para a área da linguística. Os participantes da referida pesquisa, adolescentes de 12-13 e nativos dos Países Baixos, assistiam às aulas escolares em neerlandês e inglês (50% do tempo em cada língua). No começo do estudo, os pesquisadores aplicaram questionários com o intuito de verificar as diferenças individuais dos aprendizes. Dentre os fatores observados estão a motivação (com três dimensões estudadas, cada uma inquirida por meio de nove sentenças que deveriam ter valores atribuídos para cada uma), o tempo de exposição ao inglês fora da escola e o nível de proficiência em inglês. Os alunos produziram 23 textos sobre temas debatidos em aula (histórias natalinas, primeiro mês na escola e outros) entre novembro de 2015 e maio de 2016. Os textos eram avaliados por especialistas, e a média dos alunos foi, gradualmente, se elevando de 2.1 (em uma escala de 1-5) até 2.9 no final. Também houve avaliação das produções por meio de medidas analíticas como complexidade sintática e complexidade lexical. A ergodicidade foi testada observando se as trajetórias de desenvolvimento da variabilidade e ganhos de proficiência dos alunos foram similares. A hipótese era de que maiores graus de variabilidade indicam maiores ganhos de proficiência; a variabilidade foi calculada pelo coeficiente de variação, que é a dispersão dividida pela média. Os pesquisadores apresentam um gráfico contendo todas as trajetórias dos 22 alunos subdivididos em grupos pela proficiência: os indivíduos de cada grupo, já selecionados e com mais semelhanças *a priori*, não apresentavam homogeneidade entre si. Nos cálculos de coeficiente de variação e ganho de proficiência, foi encontrada significância para o fato de que a maior variabilidade durante as 23 seções de coleta indicou maior ganho de proficiência no final do experimento. Como conclusão, os autores afirmam que trabalhos estudando não apenas grupos, mas também indivíduos são importantes, uma vez que as diferenças individuais existem e podem ser, facilmente, apagadas olhando apenas para a média de um grupo. Além disso, estudos diacrônicos para um informante, especificamente, podem não ser válidos para outros informantes – as trajetórias são únicas. Os autores insistem que trabalhos ao longo do tempo oferecem informações válidas para o estudo do processo de desenvolvimento, e que é preciso ter cuidado também ao extrapolar a trajetória de desenvolvimento de um falante para o grupo. Afinal, os sistemas linguísticos podem não ser ergódicos, isto é, uniformemente variáveis no tempo.

Em um artigo que estudava variabilidade intrafalante por meio de SDC, Verspoor *et al.* (2008) apresentaram evidências da grande variabilidade de desenvolvimento de aprendizes de língua estrangeira, apresentando momentos de progresso, regresso e padrões não lineares de aprendizado. Os autores relembram importantes características de SDCs (como as habilidades, experiências e motivações individuais de cada aprendiz, além da existência de diversos subsistemas interdependentes e instáveis) e buscam em autores da área de Psicologia da educação (como Piaget e Vygotsky) a base cognitiva para definir que os conhecimentos assimilados por uma aprendiz passam por *filtros* e adaptações do próprio sistema cognitivo. Os autores analisam trabalhos prévios na área de variação no aprendizado de segundas línguas, entre eles um trabalho da década de 1970 (como o citado Cancino *et al., 1978*) que estudava as formas negativas de verbos na língua inglesa. Foi possível observar que, embora os diferentes informantes da pesquisa, nativos de espanhol de diferentes idades e que viviam em países de inglês como língua nativa por três meses, tomavam caminhos diferentes no desenvolvimento da forma gramatical para a negativa do inglês, eles passavam por certos estágios de desenvolvimento em ordens similares (os mesmos que falantes nativos do inglês, inclusive).

2.2. PERCEPÇÃO E PRODUÇÃO DOS SONS DE L2

A relação entre a Fonética e a Fonologia é tênue, e, muitas vezes, as áreas se interseccionam. Alguns modelos teóricos explicam como o desenvolvimento acontece, e por isso, apresento os modelos de *Speech Learning Model* (SLM) e *Speech Learning Model – revised (*SLM-r) que são utilizados por uma parte considerável da bibliografia da presente pesquisa. Fatores como subsistemas de línguas e bilinguismo serão mencionados diversas vezes nas seções posteriores e estão relacionados ao objetivo principal desta pesquisa. Portanto, julguei relevante adicionar uma seção com alguns trabalhos que oferecem *insights* sobre essa interação, mesmo que o foco dos trabalhos estudados não tenha sido especificamente VOT ou oclusão de consoantes.

Vários trabalhos da bibliografia utilizada no embasamento teórico da presente dissertação se estruturaram nos modelos de SLM e adotaram também um ponto de vista de SDC ou de variabilidade, por isso destaco que tais teorias podem ser complementares, e não excludentes.

É importante descrever a versão revisada do SLM, chamada de SLM-r (FLEGE; BOHN, 2021), uma vez que diversas referências utilizadas na presente pesquisa (Flege *et al.* 2006 e Flege, 2007, por exemplo) buscam subsídio no modelo SLM para apresentar dados. Flege e Bohn (2021) afirmam que a versão original do SLM focava em diferenças entre grupos de indivíduos aprendizes de uma segunda língua após um suposto período crítico de aprendizado. Já o SLM-r possui como objetivo principal entender como os sistemas fonéticos de indivíduos se reorganizarão a partir das variáveis recebidas no processo natural de desenvolvimento de L2. Algumas características como o foco em aprendizados sequenciais na L2 após a consolidação da L1 em vez do aprendizado duplo de duas línguas na infância permaneceram as mesmas em ambos os modelos. Em Flege (1995, apud FLEGE; BOHN, 2017), nota-se que o autor admite que os conhecimentos atuais de segunda língua indicam que aprender mais cedo na vida uma língua estrangeira é melhor à longo prazo. Adultos conseguem superar crianças em estágios iniciais da aquisição, mas com o passar do tempo de aprendizagem os falantes com mais tempo de contato com a língua tendem a superar os que adquirem a língua há menos tempo. O SLM

também propunha que os *inputs* fonéticos da L2 estavam sempre disponíveis para o aprendiz, permitindo também a criação de novas categorias para novos sons. O SLM-r propõe, segundo Flege e Bohn (2021), que não existem diferenças em como vogais e consoantes de L1 e L2 são aprendidos. Uma característica basilar do SLM-r é que o processo de aprendizado da fala de uma L2 é profundamente moldado por como o sistema fonético da L1 a induz.

Flege *et al.* (2006) estudam a interação entre L1 e L2 por falantes nativos coreanos aprendendo inglês e afirmam que a L2 também exerce influência na língua materna, tornando o processo de influência L1-L2 uma via de mão dupla. Os autores separam os imigrantes coreanos em grupos pela idade em que chegaram aos Estados Unidos (o informante identificado como AO3, por exemplo, chegou aos EUA aos três anos de idade). Após trazer um compilado de pesquisas realizadas nos anos anteriores, os autores questionam o paradigma da idade crítica como único fator decisivo para a pronúncia em L2, afirmando que não se pode concluir que a única razão de um imigrante produzir os sons da L2 com sotaque é devido a uma idade avançada de imigração, uma vez que diversos adultos não apresentavam sotaque forte em uma faixa em que se esperava sotaque (FLEGE *et al., 2006*). Partindo do referencial teórico do SLM, que desenvolve a tese de que o falante cria categorias para os sons da fala, Flege *et al.* (2006) encontram que os falantes coreanos mais novos criavam as categorias das oclusivas no inglês com mais facilidade (estando menos expostos às confusões entre as categorias de L1 e L2).

Flege *et al. (*2002) sustentam a tese que os subsistemas de línguas estrangeiras interagem. Os autores investigam os processos de assimilação e dissimilação de imigrantes italianos no Canadá aprendendo inglês, por meio do estudo da vogal /e/. Um exemplo de dissimilação seria o fato de que *early bilinguals* (pessoas que adquiriram o inglês mais cedo na vida) exageravam no movimento da língua para produzir o /e/, até mais do que os falantes nativos. Os *late bilinguals (*pessoas que adquiriram o inglês mais tarde na vida, depois da infância), por sua vez, aparentavam muito menos o movimento da língua ao produzirem /e/. Flege *et al. (*2003) postulam que os efeitos da idade na aquisição podem ser observados também nas consoantes. Os autores apresentam três explicações para o porquê dessa grande diferença entre falantes adquirindo a língua em momentos distintos. Uma primeira explicação, mais gerativista, justifica que os mecanismos de aquisição da L1 ficam cada vez menos efetivos com o tempo pelo amadurecimento do corpo (haveria, portanto, um período crítico de aquisição). A segunda explicação se funda na qualidade dos *inputs* que um *late bilingual* recebe, que seriam *de pior qualidade* (palavras dos autores) para imigrantes mais velhos (ou seja, entram menos em contato com a L2, o que dificulta o desenvolvimento da língua). Os autores exemplificam que crianças, geralmente, são matriculadas em escolas sempre utilizam a segunda língua e as chances desses imigrantes, futuramente, se casarem com falantes nativos é maior (tanto a escola quanto o casamento aumentam a quantidade de *input* da L2 que os imigrantes recebem, tornando esses *inputs* mais valiosos). Os autores também comentam que, nos EUA, estudantes chineses que passaram um longo prazo no país estrangeiro (sete anos, nesse caso) apresentaram resultados de identificação superiores aos intercambistas que passaram quantidades de tempo menores (dois anos, nesse caso). A terceira explicação oferecida para o efeito da idade na aquisição se refere à interação linguística e, para que uma *fala perfeita (*por um padrão nativo, considero pelo autor, nesse caso) de segunda língua existisse, os subsistemas linguísticos não poderiam, de forma alguma, interagir. O SLM, organizado por Flege (2007, 2002) e aliados teóricos, propõe que a capacidade de um indivíduo adquirir a fala sempre permanece aberta ao

aprendizado devido à plasticidade do cérebro. Os autores também afirmam que vogais e consoantes próximas entre L2 e L1 podem ser estabelecidas em categorias fonológicas diferentes (quanto mais cedo, mais fácil esse processo ocorre).

De acordo com Martins (2016), partindo de um trabalho que investigou falantes bilíngues de Hunsrückisch, uma língua trazida por imigrantes alemães, e português brasileiro do Rio Grande do Sul, o uso constante das duas línguas acaba por gerar interferências, incidentes, principalmente no PB, por esta ser a língua usada em situações sociais (como escola). Sob uma perspectiva conexionista, as adaptações realizadas no PB eram vistas antes como estratégias de aprendizado do que como erros a partir do conceito de entrincheiramento, que define que uma habilidade se torna mais entrincheirada, portanto, mais difícil de ser alterada posteriormente, de acordo com a maior generalização de conhecimentos da L1 que incidem sobre a L2. Os efeitos disso podem ser percebidos por meio, principalmente, da fonologia, mas também do léxico. Martins (2016) continua seu raciocínio afirmando que o indivíduo tende a produzir os sons da segunda língua como seriam na língua materna e os sons mais difíceis de serem alterados são, justamente, os mais parecidos entre as duas línguas. É aqui que se encaixam as análises de oclusão e VOT da presente pesquisa. Tanto em PB quanto em inglês as consoantes oclusivas não vozeadas estão presentes, mas com uma diferença apenas no aspecto de aspiração, que é mais longa, com maiores valores em milissegundos de VOT, no inglês. Martins (2016) ainda lembra que produções desviantes podem ter outras origens, como a própria falta de conhecimento linguístico do falante.

Tendo apresentado definições e um breve histórico de SDCs, além de uma base de subsistemas, dedico o próximo capítulo a dissertar sobre a área da fonética forense, uma vez que alguns resultados da presente pesquisa (Capítulo 6, mais precisamente as análises qualitativas da Seção 6.2) podem colaborar para a comparação de locutores.

CAPÍTULO

03 FONÉTICA FORENSE

No presente Capítulo, procuro definir quais são os escopos teóricos da fonética forense com o intuito de ofertar alguma colaboração para a área. Para isso, definirei na Seção 3.1 aspectos do que é fonética e o que distingue a área da fonética acústica, além do que é fonética forense e onde se encaixa a comparação de locutores na Seção 3.2.

3.1. FONÉTICA E FONÉTICA ACÚSTICA

De acordo com Cristófaro-Silva (2003), a Fonética consiste na descrição dos sons da fala. Esta descrição pode ser feita sob várias perspectivas: a fonética articulatória estuda os mecanismos biológicos e morfológicos e permitem que os seres humanos exerçam a capacidade de fala; a fonética acústica analisa o dado físico da fala produzida, geralmente, com a ajuda de *softwares;* a fonética auditiva trata das questões perceptuais do ouvinte; a instrumental, por fim, analisa a fala com instrumentos laboratoriais. O foco da presente pesquisa é a fonética acústica, que visa analisar as formas de um registro de fala por meio de um *software*. Ao introduzir a fonética acústica, Ladefoged e Johnson (2011) afirmam que sons da fala podem ser diferenciados por três características: (1) tom, (2) intensidade e (3) qualidade[5]. Os autores exemplificam que as vogais das palavras *bad* e *bed* podem ter o mesmo tom (em escala musical) e a mesma intensidade sonora, mas variam na qualidade vocálica (a primeira vogal é produzida com a mandíbula mais aberta, resultando em alterações de medidas acústicas como duração e padrão formântico). Vogais parecidas, tanto em *bad* como em *bed,* podem ser ditas de forma mais grave ou aguda, ou também com mais ou menos intensidade. Ladefoged e Johnson (2011) também explicam as diferenças entre sons de acordo com o estado dos dois músculos localizados na laringe chamados de pregas vocais. Ao expelir o ar do pulmão para produzir a fala, estes músculos podem estar ajustados de forma a deixar o ar passar por um espaço estreito, fazendo com que haja vibração das pregas vocais e que o som seja *vozeado*. Quando as pregas vocais estão separadas e o ar passa sem causar vibração, os sons são classificados como *não vozeados*. Todas as vogais se encontram na primeira categoria, de sons vozeados. Algumas consoantes também se encaixam nessa definição, como /b/. A consoante /p/ estudada na presente pesquisa se encaixa na classificação de não vozeada, sendo distintiva de /b/, *a priori*, pela ausência de vibração nas pregas vocais. Em suma, segundo Ladefoged e Johnson (2011), os principais mecanismos articulatórios da fala são o processo de entrada/saída do ar dos pulmões, o vozeamento (ou não) do segmento, a saída (ou não) de ar pelo trato nasal e a articulação dos segmentos da cavidade oral.

A importância da Fonética e da Fonologia vai além do simples entendimento de como o sistema de fala funciona. Segundo Cristófaro-Silva (2003), a Fonética e a Fonologia possuem grande importância dentro da Linguística, uma vez que são capazes de serem

[5] Original: (1) *pitch,* (2) *loudness,* (3) *quality.* Todas as traduções de citações originalmente em inglês deste trabalho foram feitas pelo autor da presente pesquisa.

estudadas com os demais componentes da gramática, como a morfologia e a sintaxe. Como exemplo dessa conexão, em minha própria experiência de TCC, encontrada em Bernardo-Silva (2019), precisei atentar aos aspectos fonológicos e morfológicos da regra de plural das línguas portuguesa e inglesa para analisar, acusticamente, as fricativas alveolares em coda. Cristófaro-Silva (2003) também pontua que as áreas de fonética e fonologia são essenciais aos estudos de aquisição e ensino de línguas maternas e estrangeiras, uma vez que é necessário estudar como *é* (além de como se ensinar) a estrutura acústica da língua que está sendo adquirida. A autora também cita áreas importantes como a tradução, a dramaturgia e a interpretação. Às vezes, muito sentido pode se perder caso o aspecto sonoro da língua seja negligenciado em um trabalho de tradução. Além disso, existem centenas de línguas indígenas que, ainda, não foram estudadas em nosso país e, a cada língua e sistema estudado, aumentam nossos entendimentos sobre como as línguas naturais funcionam. Para mencionar mais algumas áreas de importância citadas pela autora, estão a zoobiologia, telecomunicação, linguística computacional, linguagens de sinal, fonoaudiologia e a linguística forense. Na Seção 3.2, abordarei de forma introdutória quais são as características basilares da fonética forense, passando por tópicos como a verificação de locutor e identificação de locutor, comparação de locutor e análise forense perceptiva e instrumental.

3.2. FONÉTICA FORENSE E COMPARAÇÃO DE LOCUTORES

De acordo com Cristófaro-Silva (2003), cada indivíduo apresenta em sua fala peculiaridades que o individualizam. Partindo desse pressuposto, centenas de trabalhos vêm sendo desenvolvidos nas últimas décadas com a finalidade de encontrar quais são os parâmetros individualizantes da fala. A IAFPA (Associação Internacional de Fonética Forense e Acústica) foi criada em 1991, com o objetivo de compartilhar conhecimentos e aprofundar análises técnicas no ramo. Hollien (2002) problematiza a questão da continuidade da fala de um indivíduo, pontuando que o trato vocal de um locutor é composto de órgãos cuja função primária não é falar (cavidade nasal serve ao sistema respiratório e boca ao sistema digestivo, por exemplo), estando passíveis de sofrerem alterações que podem alterar a fala com o tempo. Hollien (2002) destaca diversos estudos que apontam que estresse, emoções e a hora do dia também podem alterar a fala. Nolan (2001) também salienta que a própria linguagem pode mudar com o passar do tempo: novos códigos, sotaques, vocabulário, essas questões podem alterar a forma com que um indivíduo fala. Esses fatores demonstram o quão complexa a questão da fonética forense pode ser.

Nolan (1999, 2001) descreve duas formas de subdividir a fonética forense. A primeira, que também é descrita em Hollien (2002), está repartida entre *Speaker Recognition*[6] *(SR), Speaker Verification (SV)* e *Speaker Identification (SPID)*, e a segunda entre *Naive Speaker Recognition* e *Technical Speaker Recognition.* De acordo com Nolan (1999), o SR é um termo mais amplo e genérico para abarcar a SV e a SPID. A aplicação da SV se dá em casos em que, comumente, o locutor/falante **quer** (grifo meu) ser identificado, como em caso de reconhecimentos para transações bancárias, tomadas de decisão importantes feitas por áudio, acesso à propriedade por voz e outros. Geralmente,

[6]Equivalência dos termos: *Speaker Recognition* / Reconhecimento de locutor; *Speaker verification* / verificação de locutor; *Speaker Identification* / Identificação locutor ; *naive speaker recognition* / Análise perceptivo-auditiva (GONÇALVES, 2013) ; *Technical Speaker recognition* / Reconhecimento acústico de locutor (GONÇALVES, 2013).

a voz do locutor é colocada em teste pareada a um banco de dados existente que permite que operações ou ações sejam tomadas em virtude de a voz ser verificada com sucesso. Hollien (2002) ressalta que a pergunta "Quem está falando?" é mais importante que a pergunta "O que está sendo dito?", em casos de SV. A SPID é tida como mais complexa (NOLAN, 1999, 2001; HOLLIEN, 2002). Os casos que requerem SPID estão, em geral, associados a crimes cometidos com evidência em áudio (como sequestros, chantagens e ameaças) ou circunstâncias em que um réu possui evidência acústica que pode, ou não, incriminá-lo. Nolan (1999) explica que disfarces de voz podem ocorrer em situações de SPID, dificultando a tarefa do perito. Hollien (2002) afirma que os ruídos e a baixa qualidade de gravação também podem atrapalhar uma identificação. Outra dificuldade dramática da área é que, *a priori*, a voz precisa ser identificada entre todas as outras vozes do planeta, e não em comparação a um grupo fechado de possibilidades (HOLLIEN, 2002).

A segunda subdivisão explicada por Nolan (1997 e 1999), entre *Naive Speaker Recognition* e *Technical Speaker Recognition,* pode ser explicada da seguinte maneira: a primeira tende a contar apenas com os ouvidos de testemunhas ou pessoas sem conhecimento técnico de fonética que precisam reconhecer a voz. Já a *Technical Spearker Recognition* se utiliza de *softwares* computacionais e técnicas de processamento de sinais para ser efetuada. Nolan (1999) afirma que a combinação de ambas as técnicas, em geral, é mais segura do que o uso de apenas uma delas. Também em relação à segunda subdivisão proposta por Nolan (1997; 1999), Gonçalves (2013), ao chamar a primeira de análise perceptivo-auditiva, informa que são utilizadas como base características pessoais referentes ao sexo, condição sociocultural e econômica, presença de desvios fonéticos e outras peculiaridades. Para a segunda subdivisão, que Gonçalves (2013) nomeia em português de acústica ou instrumental, medidas físicas da fala são extraídas por *softwares,* como os parâmetros da presente pesquisa (duração de oclusão, de VOT e duração total da consoante).

Gonçalves (2013) divide as principais atividades periciais em fonética forense entre a definição da origem de determinada amostra linguística (determinar quem é o locutor), verificação da autenticidade do conteúdo gravado e depreensão do texto que foi gravado originalmente (quando o conteúdo do áudio está inteligível). Gonçalves (2013) afirma que, no Brasil, a maioria das tarefas forenses do âmbito do reconhecimento de locutor (SR) estariam mais relacionadas à verificação de locutor (SV), que, geralmente, acontece após uma amostra de áudio ser obtida (como, por exemplo, por meio da interceptação telefônica) e outra amostra, com fins comparativos, ser coletada por peritos[7]. A nomenclatura adotada pela autora, corroborada na literatura, segundo ela, a partir de 2002, é *Comparação de Locutor,* abarcando, dessa forma, SV e SPID. Na comparação de locutor, portanto, comparam-se de forma contrastiva amostras de fala (GONÇALVES, 2013) para o reconhecimento de um locutor.

O foco forense do trabalho será dado aos estudos de comparação de locutores. Identificar se os parâmetros acústicos de certo falante podem ser considerados em júri é importante, pois pode ajudar a inocentar ou culpar um réu em determinada ação. Foulkes *et al.* (2010) discutem a intersecção das áreas da Sociolinguística e da Fonética, o que pode ser, diretamente, aplicado às análises de fonética forense. Um exemplo oferecido pelos autores é o /r/ pós-vocálico em coda do inglês, pois notou-se, em um conjunto de pesquisas, que classes sociais mais privilegiadas tenderam a marcar mais o traço rótico

[7]Gonçalves (2013) informa, em sua pesquisa, que após a gravação de fala espontânea dos suspeitos, há leitura de sentenças-veículo previamente estabelecidas, o que ajuda a validar pesquisas como esta.

da consoante. Concluiu-se que a demarcação rótica do /r/ em coda estaria associada à classe social. Segundo Foulkes *et al.* (2010), os trabalhos de autores como Labov focaram em questões como classe social, sexo e idade (no início das pesquisas sociolinguísticas). Outro exemplo oferecido pelos autores foi o da variável sociolinguística (ng), produzida, entre outros contextos, no final de palavras do inglês como *anything, everything e nothing.* Notou-se que a variação entre a produção de ng com a consoante nasal velar (padrão) e a nasal alveolar (não padrão) era distribuída por classe social e sexo e não de forma aleatória. Homens tendiam a produzir a variante alveolar mais que mulheres em todas as classes sociais analisadas e a alveolar ocorria dramaticamente mais nas classes média e média baixa. Foukes *et al.* (2010) também pontuam a importância de se considerar o bilinguismo como característica sociofonética, por fatores como o aumento no inventário fonético de bilíngues e a alteração nas formas de produzir sons em ambas as línguas.

A análise de parâmetros acústicos da fala oferece recursos para a comparação forense de locutores e pode ser feita de diversas maneiras. Os parágrafos descritos a seguir se propuseram a alguns desses parâmetros. Skarnitzl e Vanková (2017) analisaram 100 homens tchecos para averiguar qual seria o valor padrão para frequência fundamental (F0) na língua tcheca, tanto em fala espontânea, como em leitura guiada. Os valores de média obtidos foram: 117,3hz (média para fala espontânea), 129,3 Hz (média para leitura), 113,1 Hz (mediana para fala espontânea) e 125,8 Hz (mediana para leitura). Os autores concluíram que o espectro de frequência para F0 na fala espontânea costuma ser menor do que na leitura de textos.

Jessen *et al.* (2005) analisaram a F0 de mais de 100 homens e como o disfarce de voz para um tom mais agudo influenciava na produção de F0. A conclusão foi que era possível observar o quanto cada falante alterava sua frequência fundamental a partir de cada elevação de frequência.

Em um trabalho que contava com 50 participantes para produzir certos disfarces de voz, Gomes *et al. (*2016) compararam valores de F0, F1 e F2 entre homens e mulheres. Os participantes leram um texto de 75 palavras em vozes normais e em tentativas de disfarces e foram realizadas análises de percepção e análises acústicas das sete vogais orais do português brasileiro. Procurou-se saber o efeito que disfarces vocálicos como protrusão labial e abertura da mandíbula possuem na produção de vogais. As autoras confirmaram tendências universais como vogais mais baixas são mais longas que vogais mais altas e outras tendências como maiores valores de F0, F1 e F2 por mulheres e maior espaço vocálico destas em comparação com homens.

Bernardo-Silva e Gomes (2019) analisaram como os ditongos e fricativas poderiam contribuir para a identificação de falantes, analisando a duração relativa dos ditongos e fricativas e cruzando os dados entre os grupos bilíngues de falantes nativas de português brasileiro e do inglês estadunidense. Concluiu-se que a duração e o vozeamento (interligados em uma relação inversamente proporcional, mais vozeamento, menos duração) podem ser parâmetros promissores para averiguar a nacionalidade dos falantes, ao se comparar brasileiros e estadunidenses.

Debruçando-se sobre a questão específica do VOT e do reconhecimento de locutor, Cheung e Wee (2008), em um artigo cujo título já carrega a indagação referente à legitimidade do VOT para fins forenses, analisam as variações de aspiração entre falantes bilíngues de cantonês e inglês. Os valores de duração do VOT do cantonês para distinção de vozeamento é mais alto que os da língua inglesa, portanto sons vozeados em cantonês poderiam soar como não vozeados em inglês. Após analisar monossílabos de cinco

falantes, três homens e duas mulheres, residentes de Hong Kong, os autores concluem que a análise de VOT como parâmetro de identificação de falantes não é forte o suficiente para o cantonês e inglês de Hong Kong. Porém, afirmam que em algum outro par de línguas a conclusão poderia ser diferente (o VOT poderia ser um bom parâmetro, portanto).

Além das questões mencionadas anteriormente, a própria área da forense abre possibilidades que não são abordadas aqui, como a *Determinação de Conteúdo*[8] e a *Autenticação de Gravação,* ambas trazidas por Nolan (1999). A primeira está relacionada a saber **o que** (grifo meu) foi dito, especificamente, em um áudio, e não quem é o locutor. Dependendo da situação, uma palavra pode designar a intenção, ou não, de uma ação. A última está relacionada a definir se uma amostra de áudio foi editada e tende a ser uma tarefa cada vez mais difícil na medida que o processo de edição fica mais moderno.

Por ser uma pesquisa em fonética acústica, o presente trabalho se propõe a analisar em qualidade (não somente quantidade) os dados acústicos das produções da consoante-alvo /p/, em detrimento de se atentar às formas morfológicas do trato vocal dos falantes (articulação), como um ouvinte percebe esses sons (percepção) ou instrumentalizar a análise da produção por meio de máquinas como raios- x. No Capítulo 2, apresentei a visão de língua como Sistema Dinâmico Complexo. Julgo que essa visão, sendo uma das mais recentes e emergentes dentro da Linguística, pode contribuir por não, simplesmente, descartar e rotular como inválidas produções desviantes do que se espera de um falante, seja ele nativo ou aprendiz de língua estrangeira. Se, por um lado, espera-se que as produções possuam uma variabilidade intrafalante pequena e com aspectos identificáveis, os SDCs averiguam as possibilidades do porquê variações de qualquer natureza existem da forma que existem.

No Capítulo 4, apresento as descrições teóricas de consoantes oclusivas, contendo temas importantes como período de oclusão e VOT.

[8] *Content Determination* e *Tape Authentication,* respectivamente, no original.

CAPÍTULO

04 CONSOANTES OCLUSIVAS

Neste capítulo, defino articulatória e acusticamente as consoantes oclusivas (Seção 4.1), os estudos sobre o período de oclusão e VOT (Seção 4.2), em contextos de uso monolíngues e bilíngues (Seção 4.3).

4.1. PARÂMETROS ARTICULATÓRIOS E ACÚSTICOS DE OCLUSIVAS

De um ponto de vista articulatório, as consoantes oclusivas (ou plosivas) são aquelas cujo fluxo de ar proveniente dos pulmões é impedido de passar, temporariamente, por algum articulador, sendo, posteriormente, realizado (solto) com uma pequena explosão (*burst)* (LADEFOGED; JOHNSON, 2011). Esses articuladores podem ser diversos: lábios (formando, por exemplo, oclusivas bilabiais), ponta da língua (como oclusivas alveolares) ou dorso da língua (como oclusivas velares), apenas para citar alguns. Em um trabalho exploratório que antecedeu o estudo principal da presente dissertação, observei brevemente o trio de oclusivas orais não vozeadas do português e do inglês /p, t, k/, porém a oclusiva escolhida para ser analisada neste trabalho é a bilabial não vozeada, /p/, presente em todos os dialetos de português e inglês.

É importante distinguir entre oclusivas orais e nasais, pois, segundo Ladefoged e Johnson (2011), as oclusivas nasais são caracterizadas por, além da oclusão, o fato de o véu palatino estar abaixado e permitir que o ar escape pela cavidade nasal, formando sons como /m, n/.

Conforme descrevem Kent e Read (2011), é importante definir os conceitos de sons vozeados e não vozeados, além de aspirados e não aspirados. Os primeiros se referem ao estado das pregas vocais: caso elas estejam vibrando quando o som for articulado, chamamos o som de vozeado. Caso contrário, ele pode ser definido como não vozeado. Em uma análise acústica, os sons vozeados possuem uma barra escura (chamada barra de vozeamento) na parte inferior de um espectrograma, enquanto os sons não vozeados não contêm essa barra ou ela se apresenta esmaecida. Os sons aspirados, por sua vez, podem ser explicados como detentores de um período ruidoso não vozeado após a soltura de um segmento oclusivo, chamando atenção, principalmente, para /p, t, k/. Schwartzhaupt (2015) traz que o intervalo de tempo da oclusiva pode ser subdividido em três partes: a total obstrução de ar produzida pelas consoantes oclusivas, a soltura de ar após a oclusão (sendo o VOT este período até o próximo) e, por fim, o começo da vibração das pregas vocais do som subsequente.

Ladefoged e Johnson (2011), ao descreverem as plosivas do inglês, imediatamente apresentam que a diferença entre /p/ e /b/ nessa língua não pode ser, exclusivamente, caracterizada pelo vozeamento, uma vez que "Ambas consoantes oclusivas são essencialmente desvozeadas[9]" (LADEFOGED; JOHNSON, 2011, *p.57*). Os autores indicam,

[9] "*Both stop consonants are essentially voiceless*"

portanto, que a aspiração é uma característica importante para caracterizar a articulação da plosiva não vozeada. Os autores também apresentam as diferenças entre as alveolares /t/ e /d/ e demonstram que o *burst* costuma ser menor em uma contraparte vozeada.

De acordo com Kupske (2016), tanto no português brasileiro (PB) quanto no inglês as produções de /p/ são articuladas da mesma forma: uma oclusão total do trato vocálico pela junção dos lábios inferiores e superiores, sem a vibração das pregas vocais, seguida da explosão e início da vogal seguinte, e consequente vibração das pregas vocais (esses dois últimos passos ocorrem também em /t, k/). A consoante /t/ é articulada de forma que a ponta da língua interrompa a passagem de ar na região dos alvéolos. /k/ é articulada por meio da obstrução da passagem de ar pelo dorso da língua na região do véu palatino. O autor também nota que no inglês existe a variação alofônica da parada glotal em alguns dialetos (como na palavra *mountain*) e adiciono que em muitos dialetos do português existe também a variação alofônica da africação em certos casos (como o /t/ antes da vogal /i/ que é africado em muitos sotaques, como o carioca e o paranaense, em palavras como *tia* e *Tiago*).

De acordo com Cristófaro-Silva *et al.* (2019), as consoantes oclusivas podem ser explicadas, acusticamente, por seis características:

- Ausência de energia;
- Barra de vozeamento;
- Ruído transiente ou soltura da oclusão;
- VOT;
- *Loci* acústico consonantal ou F2 de transição;
- Configuração espectral da soltura da oclusão.

A ausência de energia é representada, quando analisada pelo espectrograma, por um espaço limpo que apresenta queda na intensidade das formas de onda. O contato do articulador com o ponto de articulação e a oclusão total da passagem de ar gera este primeiro momento. No caso da consoante estudada na presente pesquisa, /p/, a barra de vozeamento, localizada na parte inferior e horizontal do espectrograma, encontra-se esmaecida e começa a escurecer, ou aparecer, quando o som vocálico seguinte começa. Caso a barra de vozeamento estivesse presente antes do início da vogal, a consoante oclusiva seria classificada como vozeada. A soltura da oclusão é "representada na forma de onda por estrias estreitas e verticais que são exibidas logo após a ausência de energia no sinal acústico" (CRISTÓFARO-SILVA *et al*., 2019*, p.153).* O VOT, que será explicado com mais atenção nas próximas seções, corresponde ao tempo entre a soltura e o início da vibração das pregas vocais (acusticamente início da barra de vozeamento). Um valor positivo de VOT implica em consoantes não vozeadas e VOT negativo indica que o vozeamento começou antes da articulação da vogal seguinte, ou seja, ela é uma consoante vozeada. O F2 (segundo formante) de transição está relacionado de forma inversamente proporcional com o tamanho cavidade posterior à constrição (maior F2, menor cavidade). A medida é obtida no momento de transição entre a consoante oclusiva e a vogal seguinte. Por fim, também seguindo Cristófaro-Silva *et al.* (2019), a configuração espectral da soltura da oclusão está relacionada aos maiores valores de frequência no momento específico da soltura. É importante ressaltar que a primeira característica (ausência de energia) ajuda a discriminar consoantes oclusivas das demais consoantes; o ponto b (barra de vozeamento) distingue oclusivas vozeadas das não vozeadas; o ponto c também

pode ser útil para diferenciar segmentos vozeados (menor ruído transiente) de não vozeados; os pontos d, e, e f ajudam a diferenciar o ponto de oclusão das consoantes oclusivas analisadas.

Barbosa e Madureira (2015) definem as consoantes oclusivas e africadas como "produzidas com obstrução total, formada por contato entre os articuladores, à passagem de corrente de ar. Após certo tempo /.../ a obstrução é liberada pelo afastamento dos articuladores" (p. 314). O momento correspondente à liberação dos articuladores gera um ruído transiente (explosivo para a audição, por isso o nome de *plosiva*, auditivo, e o termo oclusivo, referente à produção). Os autores também apresentam as fases de oclusão e soltura por meio de um exemplo no espectrograma.

Na Seção 6.2, proponho realizar uma análise qualitativa do período de oclusão e do VOT.

4.2. DEFINIÇÃO DE OCLUSÃO E VOT

Apresento, nas seções subsequentes, as definições empíricas do período de oclusão (4.2.1) e do VOT (4.2.2).

4.2.1. OCLUSÃO

Começo notando que diversos trabalhos que são apresentados nesta seção estudaram várias consoantes, contextos, idade de informantes e vogais posteriores à oclusiva. Porém, para priorizar comparações entre o trabalho atual e a bibliografia, o foco da revisão bibliográfica será dado, principalmente, aos seguintes contextos: falantes de português ou inglês; oclusões de /p/ em posição silábica inicial (ataque); /p/ preferencialmente seguidos de vogais /a/ (uma vez que o *glide* de *pai* e *pie*, as palavras-alvo da presente pesquisa, começa com /a/); coletadas por texto ou sentenças-veículo. Como já definido na Seção 4.1, a oclusão também pode ser denominada ausência de energia e é representada pelo espectrograma em branco em uma análise acústica (CRISTÓFARO SILVA *et al.*, 2019*)*.

Vários autores realizaram trabalhos de medição da oclusão de consoantes oclusivas. Proponho apresentar uma tabela com os principais dados quantitativos encontrados na bibliografia no que concerne ao tempo de oclusão. Os valores e os contextos fonológicos e de língua de cada autor são discutidos na sequência da Tabela 1:

TABELA 1 – DADOS QUANTITATIVOS DE DURAÇÃO DE OCLUSÃO ENCONTRADOS NA LITERATURA (MÉDIA)

Autor	Vozeada	Não vozeada	/p/	/t/	/k/
/1/ Crystal e House (1987)	50ms	50ms	-	-	-
/2/ Chen (1970, APUD Crystal; House 1987)	88ms	140ms	-	-	-
/3/ Raphael (1981, APUD Crystal; House 1987)	-	-	-	-	43ms
/4/ Luce e Charles-Luce (1985) APUD Crystal; House 1987)	68ms	86ms	-	-	-
/5/ Stathopoulos e Weismer (1983)	-	-	106ms	80ms	78ms
/6/ Lousada (2006)	88ms	143ms	155ms	146ms	108ms
/7/ Melo *et al.* (2014)	-	-	150,99ms	130,11ms	119,03ms
/8/ Berejuk e Gomes (no prelo) **	-	-	112ms	103ms	75ms
/9/ Berejuk e Gomes (no prelo) ***	-	-	112,6ms	100ms	78ms

FONTE: O autor (2021).[10]

Stathopoulos e Weismer (1983) estudaram o tempo de oclusão em consoantes plosivas por meio de seis informantes (três homens e três mulheres) e sentenças-veículo com palavras inventadas em sílabas CVCVC da língua inglesa. As oclusivas analisadas variavam conforme posição na sílaba, vozeamento e tonicidade. Um dos resultados mais importantes para a presente pesquisa encontrado pelos autores foi que as consoantes bilabiais apresentaram maiores valores de duração da oclusão do que as alveolares e velares em todos os cenários testados. A média de duração de oclusão encontrada para /p/ em posição de ataque foi de 106ms, enquanto alveolares apresentaram o valor de 80ms e velares de 78ms, mantendo o padrão da bibliografia estudada sobre oclusão.

Crystal e House (1987), em um artigo que propunha analisar de forma geral a duração de consoantes em inglês estadunidense, afirmam que uma plosiva *completa,* com todas as fases de sua articulação, não aconteciam em 100% das ocasiões na fala conectada (os autores afirmaram que encontraram 437 plosivas completas de um total de 972 identificadas em uma pesquisa anterior de 1982). É importante notar que plosivas em ataque apresentaram um índice de *completude* de 85%, ao se comparar com 33% de plosivas em coda. Em se tratando, especificamente, do período de oclusão, Crystal e House (1987) encontraram um valor de 50ms de média para o tempo de oclusão de uma plosiva em inglês estadunidense produzida por falantes nativos, não sendo a diferença entre plosivas vozeadas e não vozeadas significativa. Os autores fizeram uma divisão entre falantes *rápidos* e *lentos,* e concluíram que os falantes rápidos tendiam a produzir oclusões cerca de 5ms mais curtas que os falantes lentos. Os autores também trazem dados e valores interessantes de outras pesquisas (todas as fontes até o final do parágrafo foram retiradas de Crystal e House, 1987): Chen (1970) encontrou 88ms de duração da oclusão para oclusivas vozeadas e 140ms para oclusivas não vozeadas; Raphael (1981) encontrou 57ms de oclusão para /g/ e 43ms de oclusão para /k/; Luce e Charles-Luce (1985) encontraram 86ms para não vozeadas e 68ms para vozeadas. Com relação ao ponto de articulação, Crystal e House (1987) indicam que há aparente semelhança entre as

[10] ** Sem disfarce de voz. *** Com disfarce de voz.

oclusões de plosivas labiais, alveolares e velares, porém que as alveolares tendem a apresentar valores um pouco menores. Os autores indicam também que Fischer-Jorgensen (1964) encontrou valores maiores de duração da oclusão para bilabiais; Worth e Sakuda (1966) reportaram bilabiais não vozeadas mais longas que bilabiais vozeadas, e ambas mais curtas que dentais; Port (1979) encontrou que bilabiais teriam oclusões mais longas que velares e alveolares; dentre vários outros trabalhos que abordaram outros pontos de articulação e relações que não são essenciais para a presente pesquisa.

Lousada (2006) estudou as oclusivas do português europeu. A pesquisadora analisou a fala de seis informantes e as consoantes /p, b, t, d, k, g/ em posições de início, meio e final da sílaba, analisando diversos parâmetros como o VOT e a oclusão. A autora encontrou as seguintes médias para duração absoluta de oclusão em posição de ataque: 155ms para /p/, 146ms para /t/, 128ms para /k/, 108ms para /b/, 90ms para /d/ e 75ms para /g/, sendo a média das oclusivas não vozeadas (143ms), maior que a das vozeadas (88ms). Observa-se a tendência de o ponto de articulação interferir na duração da oclusão: bilabiais apresentaram valores médios maiores que alveolares; alveolares apresentaram valores médios mais altos que velares. Lousada (2006) também trouxe os valores encontrados para cada contexto fonológico. No caso de /pa/ em posição inicial, o valor de duração da oclusão observado foi de 138ms, com 16ms de desvio padrão. Esse valor foi, consideravelmente, menor ao se comparar com os contextos /pi/ e /pu/, que apresentaram oclusão média de 163ms (os maiores de toda pesquisa) com 28ms de desvio padrão. A média de duração de oclusão de /p/ em posição medial foi de 110ms e em posição final o valor encontrado foi 123ms.

Melo *et al.* (2014) estudaram plosivas vozeadas e não vozeadas de 11 crianças e 17 adultos falantes de português brasileiro. Os parâmetros analisados na pesquisa foram tempo de oclusão, VOT, amplitude do *burst* e duração da vogal. Os adultos escolhidos para a coleta de dados declararam todos serem monolíngues e sem histórico de bilinguismo, possuir entre 18 e 44 anos e terem dominado todo inventário fonético de português brasileiro. As crianças analisadas possuíam desenvolvimento fonológico típico e possuíam entre quatro e oito anos e 11 meses de idade. Os adultos eram estudantes universitários do primeiro período e as crianças de uma escola pública, ambos residentes de uma cidade no Rio Grande do Sul. No grupo de adultos, o valor médio da duração da oclusão em /pa/ em posição inicial foi de 150,99ms, 130,11ms para /t/ e 119,03ms para /k/, focando somente nas não vozeadas. No grupo de crianças, os valores encontrados foram 127,54ms para /p/, 125m84ms para /t/ e 117,1ms para /k/. Resultados significativos foram encontrados para as diferenças no tempo de oclusão entre não vozeadas e vozeadas (sendo as não vozeadas contendo oclusões mais longas).

Berejuk e Gomes (No prelo) mediram a duração das fases de oclusão e soltura nas consoantes /p, t, k/ produzidas por 20 informantes (10 homens e 10 mulheres), em diversos contextos fonológicos. Todas as consoantes estavam em sílabas tônicas precedidas de /a/. O banco de dados utilizado para a realização dessa pesquisa era referente a um estudo de disfarce de voz e foram consideradas a leitura de um texto com e sem disfarce de voz. Os valores médios de duração de oclusão encontrados para a fala sem disfarce foram de 112ms para /p/ (em contexto /pe/, /pu/ e /pi/), 103 ms para /t/ e 75 ms para /k/.Com disfarce, as médias de duração da oclusão para /p/ foram 112,6ms, para /t/ 100ms e para /k/ 78ms. Tais valores não apresentaram diferenças significativas (entre as falas com e sem disfarce de voz), o que evidencia o potencial de adoção desse parâmetro como individualizante em um contexto forense de comparação de locutor. Com

relação às diferenças significativas entre homens e mulheres, um resultado que recusou a hipótese nula (p=0,017) diferenciou a oclusão na produção de /pi/ sem disfarce.

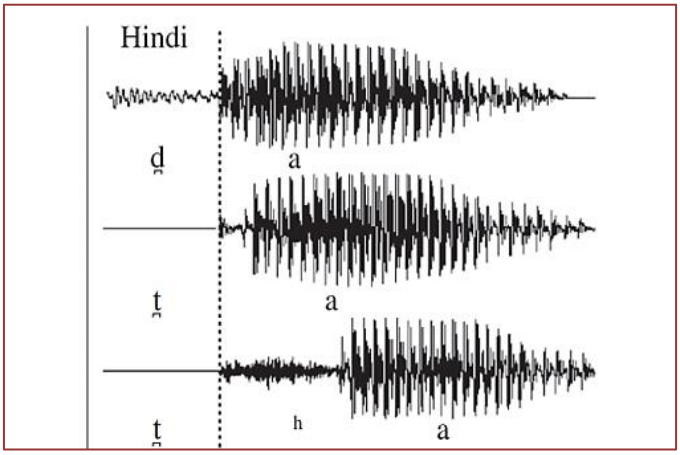
4.2.2. VOT

A definição do que é V*oice Onset Time* começou a ser traçada por Lisker e Abramson (1964), que definiram o conceito como o de tempo entre a explosão de uma consoante plosiva (a soltura do segmento oclusivo, conforme explicado na Seção 2.1) e o início de pulsação glotal (ou seja, o início da vogal seguinte). Lisker e Abramson (1964), na introdução do estudo considerado pioneiro no assunto VOT, descrevem a existência de três categorias de diferenciação das consoantes oclusivas em língua inglesa: o vozeamento, a aspiração e a força articulatória (da produção de /p, t, k/ teria mais estímulos físicos no trato vocálico na hora de serem articulados). As duas primeiras seriam observáveis, respectivamente, pelo status da glote (pregas vocais vibrando em baixas frequências para sons vozeados, ou sem a vibração das pregas vocais para os sons não vozeados) e a presença de ruído no espectrograma. A terceira categoria (força articulatória) não encontrou base empírica o suficiente para ser considerada válida pelos autores e, por isso, não tratarei dela aqui. Kupske (2016, p.76), define VOT como: "/O VOT/ /.../ faz referência ao tempo entre a soltura da oclusão e o início do vozeamento, sendo, geralmente, utilizado na produção e descrição das consoantes oclusivas". De acordo com Lofredo-Bonatto e Silva (2018), certos idiomas, como o inglês e coreano, necessitam da pista da aspiração, além do vozeamento do segmento, como traço distintivo entre sons. Tendo a pista de aspiração sendo um fator diferencial para segmentos em certas línguas pode favorecer a análise de bilinguismo de falantes de línguas em que o VOT não é um traço distintivo (como o português brasileiro, por exemplo), partindo do pressuposto que um falante pode produzir o VOT de diversas formas em sua L1 e L2. Essas variações na articulação podem ser úteis na comparação de locutores em um contexto forense.

Um VOT de duração negativa, por exemplo, ocorre quando o vozeamento do som seguinte acontece antes mesmo da soltura da consoante oclusiva. O valor da duração de VOT torna-se um número abaixo de zero, uma vez que não há tempo da aspiração acontecer. Kent e Read (2011) também afirmam que um exemplo de um VOT negativo de 130ms seria uma consoante oclusiva vozeada (como o trio contraparte de /p, t, k/, que é /b, d, g/). A classificação do que é um VOT de duração longa é controversa, porém. Segundo Reis *et al.* (2008), certos autores consideram VOTs superiores a 40ms longos, enquanto outros somente acima de 70ms. O VOT positivo é o padrão de línguas como o inglês e o sindi (KENT; READ, 2011), com valores que podem ser muito maiores que 100ms. Há também a possibilidade de haver uma produção de oclusiva não vozeada com um VOT baixo, como em alguns casos no português brasileiro (KUPSKE, 2016). Como para os nativos de português brasileiro o VOT é mínimo entre /p, t, k/, isso pode ser um fator de dificuldade no desenvolvimento da pronúncia de oclusivas no inglês.

Kent e Read (2011), trazem, na Figura 1, uma representação no que concerne às diferenças visuais em um espectrograma entre VOTs positivos, mínimos e negativos:

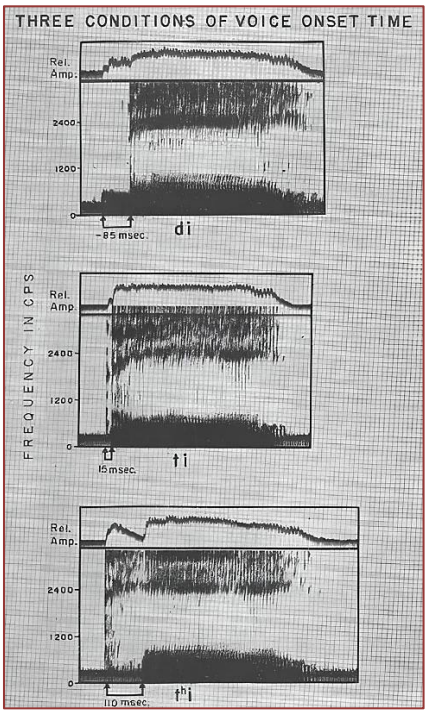
FIGURA 1 – FORMAS DE ONDAS DE VOT NO IDIOMA HINDI



FONTE: Kent e Read (2011, p.156).

De cima para baixo na Figura 1 está um VOT negativo representado pela consoante /d/, seguido por uma consoante com aspiração curta como o /t/ do português, e, por fim, uma consoante /t/ com VOT alto. O ruído na consoante /t/ com alto valor de VOT é demarcada com o diacrítico h sobrescrito à direita na transcrição fonética. Lisker e Abramson (1964), na Figura 2, também trazem um quadro similar de diferentes produções de aspiração (vozeamento prévio, pequena aspiração e muito aspirado), conforme a Figura 2:

FIGURA 2 – FORMAS DE ASPIRAÇÃO



FONTE: Lisker e Abramson (1964, p.390).

Os exemplos da Figura 2 ocorrem na mesma ordem dos exemplos da Figura 1: VOTs negativos com a consoante /d/, um VOT curto de /t/ que é, geralmente, encontrado em falantes de línguas como o português e espanhol, seguidos de um VOT alto de /t/ de línguas como o inglês. Justifica-se, portanto, o interesse em pesquisas nessa área da fonética com afirmações como as de Lofredo-Bonatto e Silva (2018):

> O parâmetro acústico do *voice onset time* (VOT) expresso em milissegundos (ms) é considerado como um intervalo de tempo decisivo na percepção acurada dos sons plosivos não vozeados /p, t, k/ e dos vozeados /b, d, g/. (LOFREDO-BONATTO; SILVA, 2018, p. 3).

Segundo Cohen (2004), espera-se que as oclusivas não vozeadas do português sejam todas de VOT curto (*short lag*). O inglês, segundo o mesmo autor, possui uma exceção: o VOT só não será longo quando /p, t, k/ forem antecedidos por /s/, como em *spare*.

Cho e Ladefoged (1999) afirmam que existem duas maneiras dentro da aerodinâmica para explicar a diferença de duração de VOT entre as consoantes velares, alveolares e bilabiais. A primeira explicação é de que, partindo do pressuposto que é necessário que haja diferença de pressão na glote para que o ar flua e os sons sejam produzidos, a cavidade supraglotal na consoante velar é menor, ao se comparar com as alveolares e bilabiais. Logo, o ar que será expelido na soltura da oclusiva possui maior pressão, e justificaria os valores maiores de VOT nas velares que nas demais oclusivas. A segunda explicação é de que, como há uma massa de ar maior à frente da oclusiva velar, mais tempo levará para ela ser movida na produção do som, aumentando o valor do VOT. Cho e Ladefoged (1999) também discorrem sobre o movimento da língua, uma vez que a ponta da língua e os lábios possuem uma velocidade de movimentação maior do que a parte anterior da língua, e estariam mais distantes do eixo de movimentação do órgão (o que geraria um tempo menor de VOT). Outra explicação é relacionada ao tempo de contato dos articuladores (CHO; LADEFOGED, 1999). A oclusiva velar é produzida pelo fechamento do trato vocal promovido pelo contato entre o dorso da língua e o palato (uma área grande, ao se comparar com a superfície da ponta da língua nos alvéolos, no caso da consoante alveolar, e da dos lábios, no caso da oclusiva bilabial), o que geraria valores maiores de VOT.

Outra consideração importante que pode ser feita no sentido de formular hipóteses sobre produções de VOT se encontra em Flege *et al.* (2003), uma vez que os autores apresentam a discussão:

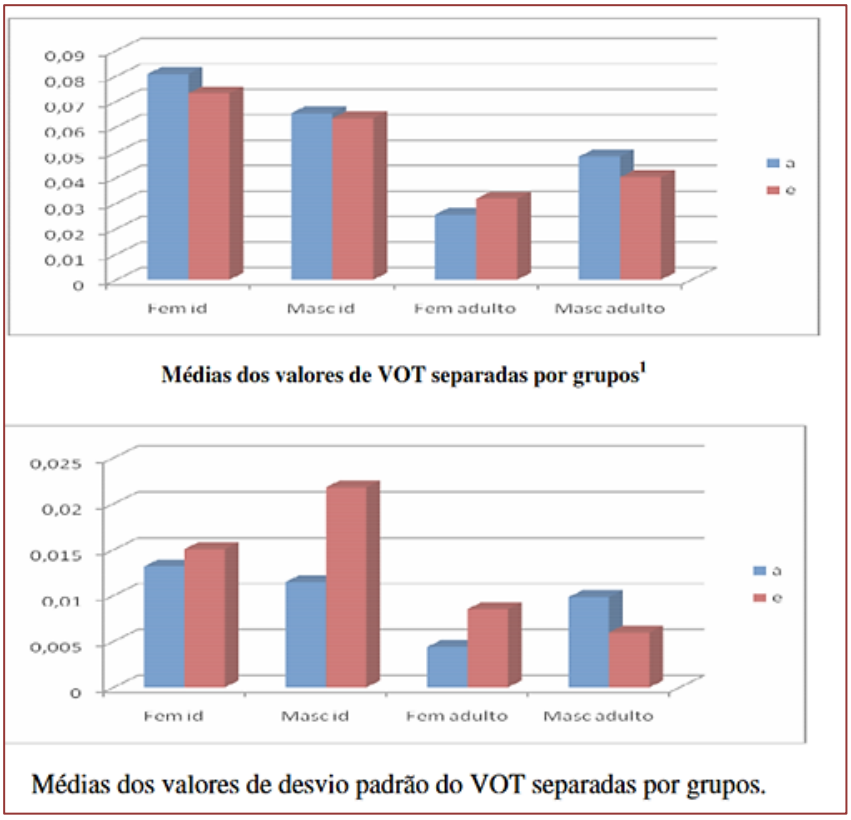
> Considere, por exemplo, um falante nativo de uma linguagem que o /t/ é produzido como uma plosiva de curta duração de VOT de cerca de 20ms (como no francês e espanhol). Este indivíduo que depois aprende uma L2 em que o /t/ é produzido como uma plosiva de grande duração de VOT, cerca de 80ms (como no inglês) pode desenvolver uma categoria misturada de /t/ que específica um valor intermediário de VOT.[11] (FLEGE *et al.*, 2003, p. 469).

[11] *Consider, for example, a native speaker of a language in which /t/ is implemented as a short-lag stop having average voice onset time (VOT) value about 20ms (e.g. French or Spanish). Such person who later learns an L2*

Na citação anterior, partindo do pressuposto do SLM, apresentado na Seção 2.2, uma segunda categoria de aspiração poderia ser estabelecida, *a priori*, como assimilação. Os autores também relatam casos práticos de valores intermediários de VOT para falantes bilíngues de francês e inglês.

Celeste e Teixeira (2009) se dedicaram a investigar efeitos de sexo, contexto vocálico e idade na produção de VOT para português brasileiro. Os informantes selecionados foram cinco mulheres adultas (23 anos de média de idade, as médias serão reportadas para todos os grupos), cinco mulheres idosas (63 anos), cinco homens adultos (25 anos) e cinco homens idosos (62 anos), nos contextos vocálicos de /a/ e /e/. Os informantes eram todos nascidos e crescidos na região de Minas Gerais, Brasil, e apresentavam escolaridade similar. O método de coleta de dados foi por meio de leitura de texto. Esperava-se que o VOT antes de /a/ fosse menor do que antes de /e/ como era descrito pela literatura, porém os resultados encontrados nesta pesquisa não foram significativos. Todavia, as autoras assumem a baixa quantidade de informantes para essa coleta e o fato dos outros trabalhos compararem /a/ com /i/, esta última sendo uma vogal ainda mais alta que /e/. A idade e o sexo dos informantes, contudo, apresentaram diferenças significativas: o grupo de adultos apresentou variações entre homens e mulheres para ambos os contextos e, no desvio padrão, como é possível observar na Figura 3, compilada de dois gráficos.

FIGURA 3 – VALORES DE DURAÇÃO DE VOT DOS GRUPOS ESTUDADOS



FONTE: Celeste e Teixeira (2009, p.34-35).

in which /t/ is implemented as a long-lag stop having an average VOT value of about 80ms (e.g. English) might develop a merged L1-L2 /t/ category that specifies an intermediate VOT value.

Além disso, o fato de os valores maiores terem sido encontrados para mulheres idosas em comparação com homens idosos, mas o oposto, entre mulheres adultas e homens adultos, reforçam a decisão de trabalhar com somente um sexo na presente pesquisa. Outro resultado interessante é a grande diferença de VOT entre idosos e adultos, os primeiros produzindo aspirações mais longas. A tentativa de trabalhar com grupos de idade bem definidos também pode ser válida.

Pesquisas importantes para o trabalho atual, como o de Kupske (2016), averiguam também produções intermediárias entre a L1 e a L2. Considero importante ter em mente que os valores encontrados poderão divergir um pouco dos valores encontrados na literatura para monolíngues e bilíngues. Por isso, a seguir, apresento alguns trabalhos revisados que estudaram o VOT em falantes bilíngues.

4.3. REVISÃO DE TRABALHOS: BILINGUISMO E VOT

A seguir, apresento um compilado dos resultados quantitativos de VOT encontrados na literatura, para facilitar a visualização. Depois da apresentação geral dos dados, discuto com mais detalhes cada trabalho encontrado na Tabela 2:

TABELA 2 – DADOS QUANTITATIVOS DE DURAÇÃO DE VOT ENCONTRADOS NA LITERATURA

Autor	Não vozeadas	/p/	/t/	/k/
/1/ Cohen (2004) Intervalo de valores PB	-	39ms-114ms	53ms-112ms	50ms-112ms
/2/ Cohen (2004) Intervalo de valores Ing	-	17ms-25ms	14ms-46ms	22ms-60ms
/3/ Alves e Seara (2008) PB	-	24,79ms	30,22ms	43,2ms
/4/ Alves e Seara (2008) Ing	-	35,61ms	46,36ms	51,13ms
/5/ Lima Jr (2015) Controle	12ms-38ms	62ms	77ms	-
/6/ Lima Jr (2015) G1	55ms-80ms	29ms	62ms	-
/7/ Lima Jr (2015) G2	-	23ms	49ms	-
/8/ Lima Jr (2015) G3	-	18ms	39ms	-
/9/ Kupske (2016) PB monolíngue	-	15,13ms	17,88ms	38,93ms
/10/ Kupske (2016) Ing monolíngue	-	56,95ms	77,31ms	82,55ms
/11/ Kupske (2016) GR0-3	-	20,75ms	30,9ms	50,21ms
/12/ Kupske (2016) GR4-6	-	27,21ms	49,02ms	70,69ks
/13/ Kupske (2016) GR8-11	-	46,56ms	63,67ms	84,18ms
/14/ Schereschewsky *et al.* (2017) monolíngues PB	-	13,19ms	-	54,46ms
/15/ Schereschewsky *et al.* (2017) bilíngues intermediários PB	-	14,03ms	-	76,56ms
/16/ Schereschewsky *et al.* (2017) bilíngues avançados PB	-	15,28ms	-	71,8ms
/17/ Schereschewsky *et al.* (2017) bilíngues intermediários Ing	-	23,2ms	-	87,82ms
/18/ Schereschewsky *et al.* (2017) bilíngues avançados Ing	-	17,96ms	-	74,29ms

FONTE: O autor[12](2021).

[12]Legenda: /6, 7, 8/ Grupos 1-3 serão descritos à frente; /11/ Kupske (2016); Grupo zero a três anos na Inglaterra; /12/ Kupske (2016); Grupo quatro a sete anos na Inglaterra; /13/ Kupske (2016).

São diversos os estudos que investigam a duração média do VOT em diferentes línguas. Cohen (2004), ao analisar brasileiros residentes no Brasil, proficientes em inglês, sendo muitas vezes professores, encontrou valores em língua inglesa entre 39ms e 114ms para /p/, 53ms e 112ms para /t/ e 50ms e 112ms para /k/. No Português Brasileiro (PB), os valores encontrados são entre 17ms e 25ms para /t/, 14ms e 46ms para /p/ e 22ms e 60ms para /k/. O trabalho de Cohen (2004) se utilizou do embasamento teórico do SLM e investigou como o VOT da segunda língua influenciava a primeira em falantes nativos de (PB) e avançados na língua inglesa, isso partindo do pressuposto que as duas línguas possuem um *status* de VOT diferente (a aspiração no PB sendo muito menor). O autor seguiu os preceitos de Flege, no sentido de esperar uma interação entre L1 e L2. A idade de aquisição também é um fator importante para ser considerado pois, segundo Cohen (2004), a idade em que o traço do vozeamento é adquirido varia entre as línguas. Isso pode ser exemplificado pela aquisição da pronúncia das oclusivas em língua inglesa estar entre um ano e 10 meses e dois anos e oito meses de idade, enquanto no espanhol estar por volta de quatro anos de idade (COHEN, 2004).

Alves e Seara (2008) estudaram a produção de VOT em português e inglês por quatro falantes nativos do PB estudantes de inglês (homens e mulheres entre 19 e 28 anos com ou sem experiência no exterior). Esperava-se que os falantes fossem produzir as consoantes do português brasileiro com leve aspiração, e no inglês com uma aspiração maior. As autoras partiram, *a priori*, da diferenciação de VOT de Lisker e Abramson (1964), descrevendo três tipos possíveis de ocorrência: pré-sonorização, retardo curto e retardo longo, sendo que as consoantes oclusivas do PB estariam nas duas primeiras categorias, enquanto as do inglês nas duas últimas. Porém, há o reconhecimento que tais fronteiras são problemáticas e elas propõem uma subdivisão em quatro contrapartes, partindo de Cho e Ladefoged (1999): não aspiradas (VOT entre 0ms e 40ms), levemente aspiradas (entre 40ms e 60ms), aspiradas (60ms <VOT<100ms) e fortemente aspiradas (VOT> 100ms, como alguns casos no coreano, por exemplo). Além disso, Alves *et al. (*2008, apud ALVES; SEARA, 2008) encontraram que algumas produções do PB poderiam ser incluídas na categoria "levemente aspiradas", com VOTs entre 40ms e 60ms. O método utilizado para coleta dos *tokens* foi leitura de textos jornalísticos em ambas as línguas. Os valores médios encontrados nas oclusivas não vozeadas do PB foram: para o participante 1, 19,3ms para /p/; 27,56ms para /t/ e 40,55 para /k/; para o participante 2, encontrou-se 26,68ms para /p/; 31,79ms para /t/ e 48,38ms para /k/ ; para o participante 3, 19,97ms para /p/; 25,01ms para /t/ e 41,44 para /k/; finalmente, para o participante 4, encontrou-se 33,21ms para /p/; 36,51ms para /t/ e 42,43 para /k/. Os valores médios para as oclusivas não vozeadas no inglês foram: para o participante 1, 24,26ms para /p/; 46,63ms para /t/ e 51,41 para /k/; para o participante 2, encontrou-se 36,29ms para /p/; 45,91ms para /t/ e 55,18ms para /k/; para o participante 3, 46,62ms para /p/; 49,78ms para /t/ e 53,17 para /k/; finalmente, para o participante 4, encontrou-se 35,26ms para /p/; 43,11ms para /t/ e 44,74 para /k/. A média total encontrada na pesquisa foi, para o PB, 24,79ms para o /p/, 30,22ms para o /t/ e 43,2ms para /k/. No inglês, os valores foram, respectivamente para /p/, /t/ e /k/, 35,61ms, 46,36ms e 51,13ms. Além da conclusão de que essa amostra de falantes foi capaz de produzir VOTs maiores em língua estrangeira, uma descoberta interessante foi não haver diferenças significativas entre aqueles que tiveram contato com inglês no exterior e aqueles que aprenderam e utilizam inglês no Brasil.

Lima Jr (2015) também compila uma tabela de valores referenciais de VOT para português e inglês. Os valores encontrados pelo autor estavam, em português, entre 12ms

e 38ms para oclusivas não vozeadas e -75ms e 0ms para oclusivas vozeadas. Para o inglês, os valores encontrados foram entre 55ms e 80ms em oclusivas não vozeadas e 0ms e 35ms para oclusivas vozeadas. Lima Jr (2015) também averigua a variável da idade na aquisição do VOT em falantes brasileiros, mas adiciona a questão do inglês como L2. O pesquisador analisou a produção de 30 brasileiros que adquiriram a língua inglesa em diferentes fases da vida (grupo 1 ou G1, antes dos 12 anos de idade, G2, entre 12 e 14 anos de idade e G3, após 16 anos de idade; a divisão por sexo foi feita em todos os grupos de forma igual, metade/metade) e comparou com um grupo controle de 10 falantes nativos estadunidenses. Partindo de uma visão de língua como SDC, o autor encontrou singularidades nos padrões de produção de cada falante, e não apenas um padrão claro de desenvolvimento da característica da aspiração. O corpus do estudo foi limitado pelas consoantes /p, t, b/[13], então não há como extrapolar os resultados das bilabiais e alveolares para as consoantes velares. Os valores médios encontrados pelo autor foram, para o grupo controle: 14ms (DP[14]=5) para /b/, 62ms (DP=23) para /p/ e 77ms (DP=20) para /t/; para o grupo 1: -57ms (DP=76) para /b/, 29ms (DP=22) para /p/ e 62ms (DP=27) para /t/; para o grupo 2: -58ms (DP=65) para /b/, 23ms (DP=21) para /p/ e 49ms (DP=29) para /t/; finamente, para o grupo 3: -90ms (DP=73) para /b/, 18ms (DP=11) para /p/ e 39ms (DP=19) para /t/. Ao analisar os dados anteriores, é possível perceber que os VOTs de /t/ foram maiores que os de /p/, que por sua vez foram maiores que os de /b/. É notável também a diminuição dos valores de VOT nos grupos mais velhos. Constatou-se por um teste ANOVA diferenças significativas entre os grupos para todas as consoantes, e por um teste *post-hoc* de Tukey que há diferenças entre todos os grupos com a exceção de G1 e G2 em /b/. O que vai além das estatísticas de grupo, porém, é o fato de que produções muito desviantes do grupo de falantes nativos não serem somente provenientes do grupo G3 (de pessoas mais velhas) – membros de todos os grupos puderam ser observados junto a diversos valores de VOT, por meio de uma plotagem de valores individuais. Como pesquisador, considero que a reflexão sobre os valores individuais e a escolha de apresentação de dados é importante para análise, uma vez que toda essa pluralidade de dados seria perdida ao olhar somente para as médias totais como uma grande porção dos trabalhos faz.

Dias *et al.* (2016) estudaram a produção de VOT no inglês por falantes brasileiros buscando a relação entre a duração da aspiração e o tempo de soltura final em palavras como *bob*. Os autores partem do pressuposto que falantes nativos realizam os segmentos /b/ da palavra de formas distintas (o último sem soltura audível) enquanto aprendizes brasileiros de inglês no nível básico iriam produzir /b/ da mesma forma em ambos os contextos. O corpus contou com 24 palavras CVC começando e terminando com os pares /p, b, t, d, k, g/. O grupo controle contou com três estadunidenses da costa leste e o de brasileiros com três estudantes universitários de uma universidade baiana. Como resultado, os autores encontraram que os falantes brasileiros produziram valores de oclusivas em coda superiores aos em *onset*, talvez pela dificuldade extra do português brasileiro não possuir plosivas em coda. Comprovou-se que os estadunidenses produziram durações relativas em *onset* maiores que em coda (média geral de 28% da palavra para *onset* e 21% para coda, enquanto a média dos brasileiros ficou 21% para *onset* e 32% para coda). Os autores também cogitaram que essa diferença poderia

[13]Ao se referir aos dados da presente pesquisa, apresento-os sempre entre colchetes, para simbolizar ocorrências fonéticas. Ao mencionar outros trabalhos, mantenho a preferência do autor pelo uso de barras ou colchetes.

[14]O Desvio-Padrão (DP) é a medida que expressa a dispersão dos valores em certo conjunto de dados.

significar que os brasileiros em nível inicial estariam transferindo características da L1 para a L2.

Kupske (2016) compila dados de VOT em português e inglês de várias pesquisas. Segundo o autor, os seguintes valores também foram encontrados em pesquisas prévias para PB. Todas os dados reportados nos pontos a seguir foram extraídos de Kupske (2016):

- Para /p/: Klein (1999), 15,59ms; Gewehr-Borella et al. (2011), 16,60ms; Reis e Nobre-Oliveira (2007) 17,27 ms (DP = 6,03 ms); França (2011); 19,56 ms (DP = 10,76 ms); M. Alves (2015) 20 ms (DP = 8 ms); Istre (1980) 12ms.
- Para /t/: Klein (1999) 16,69 ms; Figueiredo (1995) 16 ms; Reis e Nobre-Oliveira (2007) 23,55ms (DP = 8,38%); França (2011) 21,66 ms (DP = 10,21 ms); M. Alves (2015) 21 ms (DP = 7 ms); Istre (1980) 18ms.
- Para /k/: Klein (1999) 36,36 ms; Gewehr-Borella et al. (2011), 37,00ms; Reis e Nobre-Oliveira (2007) 46,55 ms (DP = 9,05 ms); França (2011), 47,20 ms (DP= 15,04 ms); M. Alves (2015) 52 ms (DP = 16 ms); Istre (1980) 38ms.

De acordo com Kupske (2016), os seguintes valores puderam ser encontrados na literatura de VOT em inglês:

- Para /p/: Toribio *et al*. (2005): 55ms; Kent e Read (2002): 46-82ms.
- Para /t/: Toribio *et al*. (2005) : 70ms ; Kent e Read (2002) : 65-95ms.
- Para /k/: Toribio *et al*. (2005): 80ms; Kent e Read (2002): 70-110ms.

Kupske (2016) estudou o processo de atrito linguístico[15] que ocorreu na fala de brasileiros vindos do Sul do País que migraram para Londres, na Inglaterra. O trabalho analisou produções das plosivas /p, t, k/. Tendo o auxílio de grupos controles de inglês britânico do Sul e português brasileiro de Porto Alegre, Rio Grande do Sul, com finalidade comparativa, o autor observa o quanto o tempo de vivência no exterior alterou a fala dos brasileiros (não somente na L2, mas principalmente na L1). O pesquisador dividiu os imigrantes por tempo fora do Brasil: entre zero e três anos, quatro e sete e oito e 11. Como resultados gerais, encontrou-se que, no inglês, os brasileiros que moravam há pouco tempo em Londres produziam todas as plosivas de formas diferentes dos nativos britânicos; os que estavam entre quatro e sete anos produziam /p/ e /t/ de forma diferente dos nativos britânicos, mas não /k/; os que moravam há muito tempo na Inglaterra produziam de forma parecida com os nativos britânicos (o que também significa valores maiores de VOT). No português, tem-se que os imigrantes recentes continuavam a produzir VOTs como nativos brasileiros; os que estavam morando entre quatro e sete anos produziam /p/ como nativos brasileiros, mas não /t/ e /k/; os que já moravam há mais de oito anos no exterior apresentaram valores superiores de VOT para todas as plosivas em comparação com os brasileiros.

Para o grupo controle de monolíngues de PB na língua portuguesa, a duração absoluta e relativa do VOT foi: 15,13ms e 1,36% para /p/, 17,88ms e 1,54% para /t/ e 38,93ms e 3,15% para /k/. Para o grupo controle de monolíngues de inglês na língua inglesa, a duração absoluta e relativa do VOT foi: 56,95ms e 5,2% para /p/, 77,31ms e

[15]Perda parcial ou total de algumas características de uma língua. O fenômeno pode ser observado de diversas maneiras, como na residência de longa data em outro país. Consultar Kupske (2016) ou Schereschewsky *et al.* (2017) para aprofundamento.

7,17% para /t/ e 82,55ms 7,8% para /k/. É importante notar que Lisker e Abramson, no clássico artigo de 1964, encontram um valor muito próximo de duração absoluta de VOT em /p/: 58ms; aproximadamente 70ms em /t/ (o que fica, consideravelmente, abaixo do valor encontrado por Kupske) e 80ms para /k/. Em relação aos dados para os grupos bilíngues: O grupo de residência em Londres de zero até três anos apresentou média de duração (absoluta em milissegundos e relativa em porcentagem da palavra) de 20,75ms/1,73% de VOT para /p/, 30,9ms/2,55% de VOT para /t/ e 50,21ms/4,32ms de VOT para /k/. O grupo de residência entre quatro e sete anos apresentou média de 27,21ms/2,35% de VOT para /p/, 49,02ms/4,09% de VOT para /t/ e 70,69ms/5,95% de VOT para /k/. O grupo entre oito e 11 anos de residência apresentou média de 46,56ms/3,86% de VOT para /p/, 63,67ms/5,36% de VOT para /t/ e 84,18ms/7,20% de VOT para /k/.

Outra importante colaboração para a presente dissertação por estudar VOTs de brasileiros bilíngues residentes no Brasil foi Schereschewsky *et al.* (2017), em um trabalho que analisou as produções de VOT em consoantes oclusivas bilabiais e velares vozeadas e não vozeadas, investigando o atrito linguístico na língua materna por brasileiros residentes no país de origem. Os 33 informantes investigados ao todo foram divididos em: 11 monolíngues de português, 11 falantes bilíngues PB/inglês intermediários e 11 bilíngues PB/inglês avançados (este último grupo corresponderia ao brasileiro médio analisado na pesquisa atual: bilíngue, mas residente no Brasil). Em palavras em português, as diferenças significativas puderam ser observadas na consoante velar surda. Partindo também de uma visão de SLM, os falantes possuem um espaço fonológico comum para seus inventários de L1 e L2, o que gera interferência mútua entre as línguas. Os valores encontrados em português foram os seguintes: 13.19ms para /p/ e 54.46ms para /k/ para monolíngues; 14.03ms para /p/ e 76.56ms para /k/ para os intermediários em inglês; 15.28ms para /p/ e 71.8ms para /k/ para os avançados em inglês. No inglês, os valores foram: 23.2ms para /p/ e 87.82ms para /k/ para os intermediários; 17.96ms para /p/ e 74.29ms para /k/ para os avançados. Ao se comparar as leituras de palavras em português e inglês, foi encontrada diferença significativa entre as línguas e entre monolíngues e bilíngues, mas não entre os níveis intermediário e avançados (em língua inglesa). Nestes, os valores de VOT para os falantes intermediários foi um pouco mais alto (SCHERESCHEWSKY *et al.*, 2017). Na conclusão, os autores mencionam que há evidências que as oclusivas velares teriam maiores valores de VOT no português falado no sul do Brasil, o que poderia ter ajudado os falantes a produzirem um VOT mais parecido com os de nativos do inglês.

Schereschewsky *et al.* (2019) estudaram o atrito linguístico em falantes nativos de português brasileiro e bilíngues de inglês como L2 ou trilíngues de inglês como L2 e alemão como L3, partindo de um pressuposto de multidirecionalidade entre as línguas de um falante. Os segmentos estudados foram oclusivas iniciais não vozeadas em ataque silábico /p, t, k/ e a duração de VOT foi o parâmetro quantificado. Os autores contaram com 40 voluntários, divididos entre grupos-controle de nativos (português, 10 pessoas de Porto Alegre (RS), e inglês, 10 pessoas de Londres) e grupos experimentais (bilíngues PB/inglês, nativos do Rio Grande do Sul e estudantes de nível superior de Inglês, e trilíngues PB/ inglês/ alemão, nativos do Rio Grande do Sul e que aprenderam inglês como L2 e alemão como terceira língua (L3). O método de coleta se efetuou meio de leitura de sentenças-veículo (18 palavras-alvo por língua), com três repetições em ordens diferentes e com a presença de distratores. Destaco maior atenção nas produções de /p/ em inglês e português. Em inglês, o grupo de nativos produziu os /p/ com uma média de

56,95ms (entre 25,96ms e 90,97ms); o grupo de bilíngues produziu /p/ com média de 40,02ms (entre 22ms e 79,4ms) e o grupo de trilíngues produziu /p/ com média de 41,72ms (entre 37,21ms e 49,3ms). Nota-se a aproximação da produção dos bilíngues e dos trilíngues, que ficou apenas 1,7ms maior para os trilíngues) e a produção mais longa dos monolíngues britânicos. Para o segmento /p/ em inglês, os trilíngues apresentaram menores desvio-padrão que os bilíngues e os próximos monolíngues. Em português, os nativos produziram /p/ com média de 15,13ms (valores entre 7,31ms e 25,09ms), enquanto os bilíngues apresentaram valor médio de 27,54ms (entre 14,67ms e 41,19ms) e os trilíngues 27,05ms de média (entre 21,81ms e 37,64ms). Assim como no inglês, as maiores diferenças residiram entre os grupos de bilíngues e trilíngues para os monolíngues do que para bilíngues e trilíngues entre si. De forma geral, observou-se que, para os trilíngues, formou-se uma categoria única para o VOT de inglês e alemão. Os bilíngues estabeleceram duas categorias para português e inglês e a categoria de inglês coincide com a de alemão para os trilíngues. A conclusão a que os autores chegaram a partir dessa observação foi que o número de categorias fonológicas criadas é mais importante do que o de línguas adquiridas por um informante.

A partir de um ponto de vista complexo/dinâmico, Schereschewsky (2021) estudou VOT em sistemas multilíngues de português como língua materna, inglês como L2 e francês como L3. Os informantes eram cinco no total e foram realizadas 12 coletas de dados em um espaço temporal de três meses, com seis intervenções pedagógicas de instrução explícita ocorrendo no período correspondente às coletas de número quatro e nove). Utilizando-se de três métodos de investigação empírica (análises de correlações móveis, análises de pico e análises de *change-points*), a pesquisadora encontrou que os períodos de melhoria de desempenho mais significativos ocorreram durante a janela de tempo em que as intervenções pedagógicas aconteciam; períodos de correlações positivas e negativas ocorriam durante o período de desenvolvimento de VOT; também ocorreu uma alta variabilidade no desenvolvimento do processo de aprendizagem, o que vai ao encontro de estudos como De Bot *et al.* (2007), Verspoor *et al.* (2008) e Lowie e Vespoor (2019), como exemplos de trabalhos que abordam a complexidade. Como a pesquisa de Schereschewsky (2021) é extensa, ressalto apenas uma parte da análise que se refere às médias de VOTs de cada informante no começo da pesquisa e suas trajetórias para a consoante /p/. A autora realizou uma média aritmética para 17ms de VOT em /p/ para falantes monolíngues nativos de PB e 63ms de VOT para /p/ para falantes monolíngues de inglês. O informante 1 apresentou os seguintes valores de /p/, na primeira coleta: 30,94ms e 20,52ms de DP (português); 25,07ms e 14,82ms de DP (inglês). A informante 2 apresentou: 29,87ms e 16,11ms de DP (português); 38,35ms e 13,37ms de DP (inglês); os valores para a informante 3 foram 39,98ms e 21,15ms de DP (português); 45,99ms e 23,35ms de DP (inglês); a informante 4 teve como médias 33,38ms e 17,62ms de DP (português); 32,02ms e 11,81ms de DP (inglês) e a informante 5 teve 36,38ms e 15,51ms de DP (português) e 53,51ms e 27,7ms de DP (inglês). Os valores de VOT em português já estavam, consideravelmente, acima dos valores descritos para monolíngues na literatura, o que poderia indicar um sinal de atrito. O informante 4, aparentemente, não diferenciava português e inglês, produzindo um VOT até um pouco maior em sua língua materna. Os informantes apresentaram médias em inglês, no geral, abaixo da literatura de monolíngues, embora algumas informantes (como a 5) tiveram médias mais próximas aos valores na casa de 60ms da literatura de /p/. Após as doze coletas e três meses, com sessões de instrução explícita ocorrendo nos meses intermediários, alguns valores de VOT mudaram dramaticamente. O informante 1, por exemplo, apresentou média de 18,82ms e 4,90ms de DP para /p/ em português, e 65,54ms e 20,72ms de DP para /p/ em inglês,

diminuindo sua média de duração de VOT em português quase pela metade do que na primeira coleta e mais que dobrando sua produção de VOT para /p/ em inglês, chegando em níveis de monolíngues da língua inglesa. A informante 2 apresentou média de 33,27ms e 15,56ms de DP para português na última coleta, aumentando um pouco seu VOT na língua materna. O VOT em inglês, por sua vez, teve como média 59,64ms e 38,04ms de DP na última coleta (cabe notar que em coletas intermediárias a média de VOT chegou a valores como 93,25ms na quarta coleta, 76,61ms na quinta coleta e 95,44ms na sétima coleta), um valor significativamente mais elevado do que no início da pesquisa. Vesrpoor *et al.* (2008), em um estudo mencionado no Capítulo 2, já haviam encontrado processos de progressão e regressão variáveis no tempo. Saliento a importância dessas variações em se tratando de trabalhos que abordam as linhas da complexidade.

A seguir, no Capítulo 5, apresento a metodologia de pesquisa do trabalho, começando pelo experimento exploratório, passando pelo estudo principal e finalizando com as hipóteses da pesquisa.

CAPÍTULO

05 METODOLOGIA

Procuro descrever aqui quais foram os procedimentos metodológicos adotados tanto para o experimento exploratório (Seção 5.1), que desenvolvi nos primeiros meses do mestrado, quanto para o estudo desta pesquisa (Seção 5.2). Na Seção 5.3, retomo os objetivos do trabalho e apresento as hipóteses da pesquisa.

O desenvolvimento da metodologia do trabalho se deu junto à realização de um experimento exploratório na metade de 2020. Com o avanço da pandemia de coronavírus e a impossibilidade de encontrar informantes, pessoalmente, em ambientes fechados como estúdios de gravação, foi necessário aceitar bancos de dados gentilmente cedidos por outros pesquisadores. Sem esse auxílio, os dados, provavelmente, precisariam ser coletados de forma remota, talvez acarretando a perda de qualidade acústica para análises mais acuradas.

Para o experimento exploratório e para a pesquisa, foram utilizados os mesmos dois bancos de dados (GOMES, 2018), um contendo um texto de seis partes (lidas uma vez em cada língua, português e inglês, por cada informante) e uma sequência de 40 sentenças-veículo (lidas três vezes por cada informante) contendo, além das palavras-alvo, distratores (vide Apêndices 1 e 2 para a lista completa). As sentenças-veículo possuíam a estrutura "*Say pies badly*" em inglês e "*Diga pais baixo*" em português, sendo repetidas três vezes por cada informante e apresentadas em ordens diferentes a cada leitura. O número total de informantes do banco de dados, entre homens e mulheres, foi de 10 estadunidenses (intercambistas ou imigrantes vivendo no Brasil, com média de 12 anos de residência no Brasil) e 22 brasileiros falantes de inglês. O instrumento utilizado para coleta de áudio foi do modelo *Zoom H1 Handy Recorder*, com dois microfones embutidos. Os dados foram coletados em um laboratório de acústica de uma universidade federal no estado do Paraná. Os informantes assinaram um termo de consentimento e realizaram um teste antes da gravação oficial, que ocorreu, individualmente, na cabine isolada com o texto e as sentenças-veículo impressas em papel, dispostas em uma mesa, que também continha o gravador. Na presente pesquisa, foram selecionadas 14 mulheres, que serão descritas na Seção 5.2.

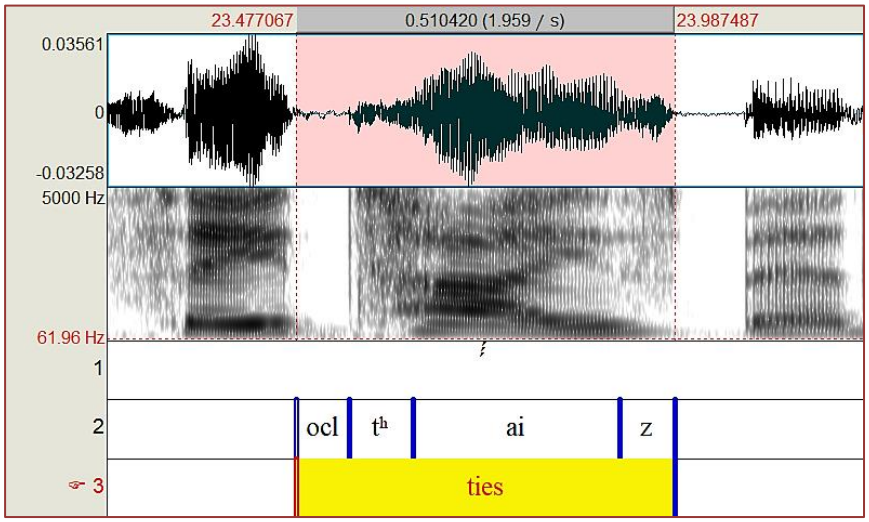
Para a análise dos dados, fiz uso de estatística descritiva (com gráficos de barra e diagramas de caixa) e inferencial tanto para o experimento exploratório como para a pesquisa. Após verificar que alguns dados não possuíam distribuição normal, foram usados testes estatísticos não-paramétricos (*Mann-Whitney* para comparações entre os grupos e testes de *Wilcoxon* para comparações intragrupos) para realizar as testagens estatísticas, tendo o valor de p igual ou menor que 0,005 como indicador de resultados significativos. Por resultados significativos se implica, basicamente, que há menos de 5% de chance de as diferenças entre os dados serem mero acaso ou aleatoriedade.

5.1. EXPERIMENTO EXPLORATÓRIO

Uma análise exploratória foi efetuada com duas informantes (uma brasileira e uma estadunidense) para que pudesse me familiarizar com os dados e as possibilidades de análise e, posteriormente, fazer um recorte para esta pesquisa. A informante B7 possuía 19 anos, era estudante de Letras Inglês e professora de inglês e residente no Brasil. A informante A4 possuía 32 anos de idade, era professora nativa de inglês e residente no Brasil. Ambas afirmaram utilizar as duas línguas, PB e inglês, com frequência. Essas informantes foram escolhidas ao acaso entre as demais participantes. No experimento exploratório, etiquetei todas as ocorrências de consoantes /p, t, k/ oclusivas em ataque, independentemente do segmento que vinha a seguir (monotongo, ditongo, tritongo e outros) nos instrumentos texto em inglês e sentenças em inglês. Todas as demarcações foram as seguintes: começo e final da palavra; começo da palavra até a soltura da plosiva, soltura da plosiva até início da vogal (VOT); início da vogal até o final da palavra (para, desta forma, ter a duração total da palavra e, assim, pode calcular as médias relativas de duração da oclusão, do VOT e da duração total da consoante, esta composta pela soma dos valores do VOT e da oclusão). Como já mencionado na introdução, relembro que, em decorrência de os VOTs em português costumarem ser muito pequenos (chegando a valores até menores que 10ms), optei por medir este parâmetro contendo o *burst* da plosiva em todos os casos.

Para a realização das análises acústicas, optei pelo software livre *Praat*[16], desenvolvido por pesquisadores/linguistas da Universidade de Amsterdam. Os valores da duração da oclusão e VOT foram medidos tanto de forma absoluta (valor total em milissegundos) quanto relativa (como forma de normalização dos dados em decorrência de possíveis diferentes velocidades de fala). Os valores de durações relativas podem ser obtidos dividindo o valor da duração absoluta de VOT pelo valor total da duração da palavra e multiplicando por 100, como em Bernardo-Silva e Gomes (2019). Na Figura 4, represento como foram feitas as marcações de palavra, oclusão, VOT e ditongos:

FIGURA 4 – EXEMPLO DE ETIQUETAGEM



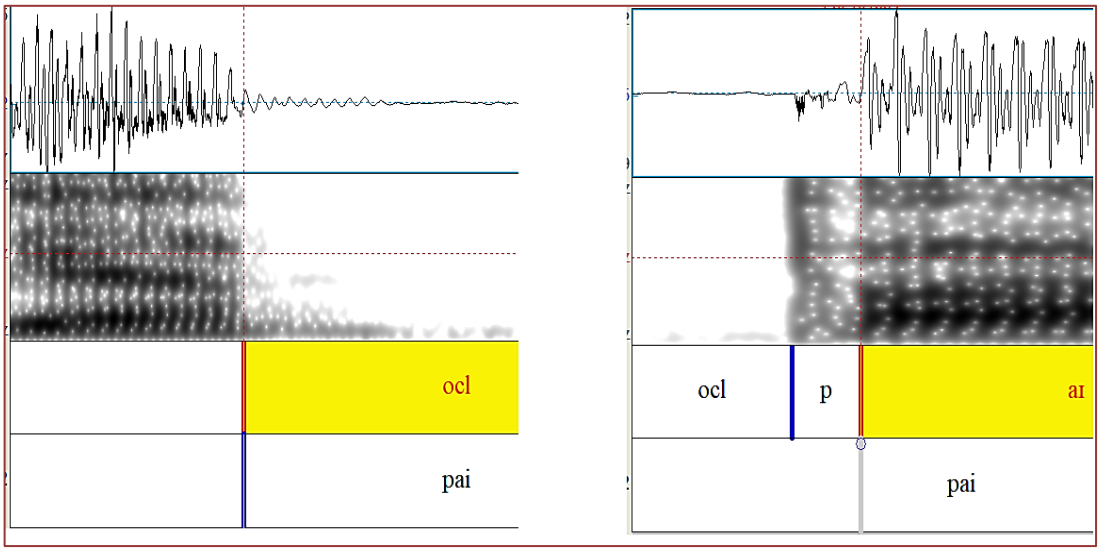
Fonte: O autor[17](2021).

[16] Identificado nas referências pelo primeiro autor, Boersma, P. Weenik, D.

[17]Apenas para fins explicativos, destaco que alguns desses dados já foram utilizados na pesquisa de Bernardo-Silva e Gomes (2019), por isso os ditongos estão subdivididos entre vogal e *glide*. Como o ditongo

A demarcação do início da palavra-alvo foi feita no espaço em branco que representa a oclusão dos articuladores (lábios superiores e inferiores se juntando) após a palavra anterior. Assim que o *burst* acontece, faço a demarcação do início do VOT. Kupske (2016) afirma que o padrão para se começar a medição de VOT é no momento da soltura da oclusiva. O início do ditongo é marcado pela primeira forma de onda vocálica e o final dele pela transição para fricativa (mudança da forma de onda para o ruído fricativo) ou pelo final da palavra em si quando não há fricativa posterior. Na Figura 5, demonstro com mais detalhes onde foram etiquetadas as fronteiras de palavra, VOT e final da plosiva.

FIGURA 5 – DETALHES DE DEMARCAÇÕES DE ETIQUETAGENS NAS TRANSIÇÕES



FONTE: O autor (2021).

Cabe notar que Abransom e Whalen (2018), ao revistar o artigo clássico de Lisker e Abransom (1964), etiquetam duas oclusões diferentes: a vozeada e a não vozeada. Porém, como meu foco incide sobre consoantes não vozeadas, optei pela demarcação de apenas um período de oclusão[18].

O trabalho exploratório foi de grande utilidade para se familiarizar com os dados e o extenso processo de etiquetagem. Porém, era necessário um filtro para que o trabalho pudesse ganhar um rumo mais definido. Assim, para a pesquisa principal, apresentada na Seção 5.2, o foco foi direcionado ao segmento /p/. O motivo da escolha foi o fato de o contexto fonológico de /p/ nas palavras *pie* e *pai* em inglês e português ser o mais próximo possível nos textos e nas sentenças-veículo, além da pronúncia das palavras ser semelhante nas duas línguas. As palavras com /t, k/ não apresentavam a mesma equivalência nas duas línguas e instrumentos, mesma quantidade de repetições ou semelhança fonética. Com a opção por *pie* e *pai,* é possível encontrar um número

não é o foco da dissertação e poderia gerar confusão, etiquetei, novamente, as amostras das análises qualitativas na Seção 6.2.

[18]Essa questão pode ser importante em pesquisas futuras, então discuto com mais detalhes o tema nas considerações finais, Capítulo 7.

suficiente de ocorrências para permitir a realização de testes estatísticos entre as nacionalidades e línguas.

5.2. O ESTUDO

Para esta pesquisa, analisei, acusticamente, a produção do par de palavras *pie* e pai, por 14 informantes no total: sete falantes brasileiras e sete estadunidenses. As falantes brasileiras eram provenientes da Região Metropolitana de Curitiba, estudantes de Letras Português/Inglês, Letras Inglês ou professoras de inglês e declararam fazer uso das línguas portuguesa e inglesa em sua rotina. As estadunidenses eram também sete, vindas da costa leste ou nordeste dos Estados Unidos, declararam uso de ambas as línguas e eram residentes da região de Curitiba no momento da coleta. Optou-se por selecionar somente informantes mulheres para controlar a variável sexo. Apresento, no Quadro 2, as informações de cada informante:

QUADRO 2 – PERFIL DAS INFORMANTES

Informante	País de origem	Região	Idade
B7	Brasil	Curitiba, PR	19
B12	Brasil	Campo Largo (Região Metropolitana de Curitiba, RMC), PR	23
B14	Brasil	Curitiba, PR	50
B15	Brasil	Araucária (RMC), PR	22
B17	Brasil	Curitiba, PR	38
B18	Brasil	Curitiba, PR	25
B20	Brasil	Curitiba, PR	26
A3	EUA	NE[19], EUA	24
A4	EUA	N, EUA	32
A5	EUA	N, EUA	34
A6	EUA	N, EUA	60
A7	EUA	NE, EUA	43
A8	EUA	NE, EUA	18
A9	EUA	-	-

FONTE: O autor (2021).

A média de idade das brasileiras (identificadas pelo prefixo B na tabela acima) analisadas foi 29 de anos, enquanto a média das estadunidenses (identificadas pelo prefixo A na tabela acima) foi de 35 anos. Todas as informantes declararam que eram fluentes e utilizavam ambas as línguas em seus respectivos cotidianos.

Depois de etiquetar as consoantes /p, t, k/ na pesquisa exploratória e perceber que a duração oclusão e a da duração total da consoante também eram parâmetros válidos de análise, optei por medir as durações da consoante /p/. Conforme mencionado no final da Seção 5.1, o par de palavras *pie*/pai foi a escolha de pesquisa, pois se apresentava disponível, diversas vezes, ao longo do texto (três em português e quatro em inglês) e das sentenças-veículo (três vezes em cada língua), permitindo uma gama maior de comparações. Os instrumentos de pesquisa (texto e sentenças) podem ser encontrados nos Apêndices 1 e 2. No banco de dados utilizado (o mesmo mencionado na Seção 5.1), essas palavras foram as que apareciam no contexto morfológico e fonológico mais

[19] NE = Nordeste dos Estados Unidos; N = Norte dos Estados Unidos.

semelhante entre as línguas. Os dados foram retirados do mesmo banco de dados mencionado, anteriormente, na Seção 5.1. Os parâmetros e etiquetagens foram os mesmos demonstrados nas Figuras 4 e 5. O trecho do texto que continha a palavra *pie* era o primeiro, e se encontra na sequência[20]:

"*A classic apple **pie** takes a shortcut with easy unroll-fill refrigerated **pie** crust. Try this sumptuous Bramley apple **pie** - simple to make and great to prepare in advance for a dinner party. This was my grandmother's apple **pie** recipe. I have never seen another one quite like it. It will always be my favorite and has won me several first-place prizes*".

O trecho para a palavra pai, no português, era:

"Como você já deve saber, no Brasil o Dia dos Pais é comemorado no segundo domingo de agosto. Esse é um dia muito especial. Você que tem um **pai** ou alguém que representa o seu **pai** na sua vida, não deixe de abraçá-lo e dizer o quanto ele é importante para você. Ter seu **pai** é um privilégio!".

Já as sentenças-veículo eram da estrutura "*Say* pie *badly"* em inglês, e "*Diga pai baixo"* em português.

O número total de palavras analisadas foi o seguinte: 14 informantes x 7 ocorrências nos textos x 6 ocorrências nas sentenças-veículo = 588. Como analisei três parâmetros (duração de oclusão, VOT e duração total da consoante) para cada palavra, o número total de dados foi 1764. Os dados em inglês estavam completos, porém houve a perda de uma repetição de "pai" nas sentenças-veículo em português pela falante identificada como B18, além do texto em português pela informante A7, diminuindo o número total para 1752.

No Capítulo 6, apresento e discuto os resultados encontrados. Na Seção 6.1, relato os resultados quantitativos da pesquisa. Apresento as estatísticas descritivas por meio de recursos como gráficos de barras, dispersão e diagramas de caixa, além de testar hipóteses por intermédio de testes estatísticos. Conforme mencionado no começo do Capítulo 5, realizei testes de normalidade entre vários conjuntos de dados e verifiquei que seria mais adequado optar por testes não paramétricos (*Mann-Whitney* para comparações internacionalidades e testes de *Wilcoxon* para comparações intranacionalidades). Considero resultados de *p* iguais ou inferiores a 0,05 como valores significativos.

5.3. HIPÓTESES

Após toda bibliografia revisada, proponho algumas hipóteses para o trabalho, que serão confirmadas ou refutadas na Seção 6.2. Retomo os objetivos da pesquisa para formular as hipóteses:

- Objetivo principal: analisar a duração de oclusão, VOT e duração total da consoante na produção de /p/ por falantes brasileiras e estadunidenses em português e inglês.
- Objetivos específicos: (a) verificar, em uma análise intergrupos, se as durações de Oclusão, VOT e duração total da consoante são diferentes em cada uma das línguas; (b) verificar, em uma análise intragrupo de estadunidenses, se as durações de Oclusão, VOT e duração total da consoante são diferentes em cada uma das línguas; (c) verificar, em uma análise intragrupos de brasileiras, se as durações de Oclusão,

[20] Vide Apêndice 1 para os textos completos.

VOT e duração total da consoante são diferentes em cada uma das línguas; (d) verificar, através da comparação de durações de oclusão, VOT e duração total da consoante produzidas pelos dois grupos nas duas línguas, se existem diferenças significativas de acordo com a coleta de dados efetuada, ou seja, por leitura de sentenças-veículo ou leitura de texto. (e) discutir quais são as implicações desta pesquisa para a fonética forense. A partir destes objetivos, quatro hipóteses foram desenvolvidas:

- (1): O grupo de estadunidenses produzirá oclusões significativamente mais curtas, VOTs significativamente mais longos e consoantes parecidas com o grupo de brasileiras em ambas as línguas.
- (2) O grupo de estadunidenses produzirá, em inglês, oclusões significativamente mais curtas, VOTs significativamente mais longos e a duração total da consoante será parecida na comparação com o português.
- (3) O grupo de brasileiras produzirá, em inglês, oclusões significativamente mais curtas, VOTs significativamente mais longos e a duração total da consoante será parecida na comparação com o português.
- (4) Haverá diferenças significativas entre as tarefas de texto e sentenças-veículo ao observar, em ambas nacionalidades e línguas, durações de oclusão, VOT e duração total da consoante.

Após revisar os objetivos e apresentar as hipóteses, apresento no Capítulo 6 os resultados encontrados.

CAPÍTULO

06 RESULTADOS

O objetivo deste capítulo é descrever os dados, que são durações absolutas, relativas, imagens de espectrograma analisadas qualitativamente e *insights* referentes à complexidade e forense, todos obtidos por meio da etiquetagem, medição e análises de duração de oclusão, VOT e duração total da consoante de 14 informantes, descritas na Seção 5.2. Este capítulo está divido em três seções. A Seção 6.1 traz os resultados quantitativos descritivos e inferenciais, a Seção 6.2 traz análises qualitativas e a Seção 6.3 revisa as hipóteses propostas na Seção 5.3.

6.1. RESULTADOS QUANTITATIVOS

Esta seção tem por objetivo apresentar a análise quantitativa dos dados por meio das médias totais de cada nacionalidade para cada parâmetro (oclusão, VOT e duração total da consoante). Os resultados de duração absoluta contam com o apoio de gráficos e os valores foram testados estatisticamente por meio de testes não paramétricos, como explanado no Capítulo 5. Os resultados de durações relativas são apresentados como forma de verificar se a velocidade de fala não é um fator que influencia nas medidas (se confirma que os resultados de durações absolutas e relativas são congruentes entre si). O *software* escolhido para as análises estatísticas foi o Jamovi e essa escolha se deu por ser um programa gratuito e fácil de usar. Para comparar, estatisticamente, o grupo de brasileiras com o grupo de estadunidenses, fiz uso do teste estatístico de *Mann-Whitney*. Para comparar a mesma nacionalidade em duas línguas, fiz uso do teste estatístico de *Wilcoxon*.

Na Tabela 3, apresento os valores totais de duração de VOT e oclusão por nacionalidade, comentando os dados posteriormente. As médias de duração absolutas estão em milissegundos e as médias de durações relativas estão em porcentagem da palavra. Em parênteses, encontram-se os respectivos desvios-padrão (DP).

TABELA 3 – MÉDIAS TOTAIS DE DURAÇÃO VOT, OCLUSÃO E DURAÇÃO TOTAL DA CONSOANTE POR NACIONALIDADE

Parâmetros / Línguas	Inglês (*pie*)	Português (pai)
Duração Absoluta VOT Estadunidenses	72,6 (6,92)	26,0 (15,4)
Duração Absoluta VOT brasileiras	47,6 (22,5)	14,9 (11,4)
Duração Relativa VOT estadunidenses	18,1 (1,46)	6,3 (2,29)
Duração Relativa VOT brasileiras	11,5 (4,82)	4,4 (3,02)
Duração Absoluta Ocl Estadunidenses	92,2 (6)	106,9 (18,9)
Duração Absoluta Ocl brasileiras	115,2 (12,1)	119,4 (21)
Duração Relativa Ocl Estadunidenses	23.3 (1,95)	26,3 (9,44)
Duração Relativa Oc brasileiras	28,0 (3,24)	32,8 (3,73)
Duração Absoluta Con Estadunidenses	164,8 (9,52)	132,9 (31,2)
Duração Absoluta Con Brasileiras	162,8 (22,3)	134,3 (16,4)
Duração Relativa Con Estadunidenses	41,8 (2,87)	32,9 (3,13)
Duração Relativa Con Brasileiras	39,5 (3,13)	37,2 (2,88)

FONTE: O autor (2021).

Com relação às oclusões como um todo, é possível retomar a pesquisa bibliográfica realizada na Subseção 4.2.1, que diz respeito aos valores de duração da oclusão encontrados na literatura. Encontra-se, em inglês: Crystal e House (1987), com 50ms para não vozeadas e vozeadas; Chen (1970), com 140ms para oclusivas não vozeadas; Luce e Charles-Luce (1985), com 86ms para não vozeadas; Stathopoulos e Weismer (1983), com /p/ em posição de ataque com 106m; Lousada (2006), com 155ms para /p/ em posição de ataque, além de 143ms de média para oclusivas não vozeadas e 138ms para /pa/ em posição inicial; Melo *et al*. (2014), com 150,99ms para /pa/ em posição inicial; Berejuk e Gomes (no prelo), com 112ms para /p/ sem disfarce de voz e 112,6ms para /p/ com disfarce. Na presente pesquisa, foram encontrados os valores de 92,2ms para estadunidenses em inglês, 115,2ms para brasileiras em inglês, 106,9ms para estadunidenses em português e 119,4ms para brasileiras em português. Os valores das brasileiras em inglês e português e das estadunidenses em português ficou próximo ao encontrado por Berejuk e Gomes (no prelo). A duração da oclusão das estadunidenses em português foi a mesma de estadunidenses em inglês de Stathopoulos (1983). O valor mais próximo para as estadunidenses em inglês foi o de 86ms encontrados por Luce e Charles-Luce (1985). Nenhum valor de duração de oclusão ficou tão alto (acima de 130ms) como em Melo *et al*. (2014), Lousada (2006) e Chen (1970), nem tão baixos (50ms) como em Crystal e House (1987).

Comparando os resultados obtidos com a base teórica estudada[21], chegamos aos seguintes resultados: o trabalho atual encontrou para o VOT de /p/ os valores de 72,6ms para estadunidenses em inglês e 14,9ms para brasileiras em português. Esses valores se encaixam, perfeitamente, na compilação feita por Lima Jr. (2015), que informa que os VOTs típicos de oclusivas não vozeadas do inglês estão entre 55ms e 80ms, enquanto os do português estão entre 12ms e 38ms. Levando em conta o estudo de Lima Jr (2015), é possível concluir que as brasileiras não produziriam um VOT suficientemente elevado para ser caracterizado como típico do inglês e as estadunidenses conseguem reduzir seu

[21] As tabelas com os dados compilados se encontram na Seção 4.3, bilinguismo e VOT.

VOT a ponto de se encaixar no típico do português. Kent e Read (2011) encontraram valores entre 46ms e 85ms para /p/ em inglês por nativos. Os valores encontrados por Cohen (2004) em língua inglesa estiveram, para /p/, entre 39ms e 114ms. Destaco que só houve seis produções de VOT acima de 100ms em todas as etiquetagens em inglês por estadunidenses, sendo metade delas por uma só informante (A9). Schereschewsky *et al.* (2017) encontraram, na língua inglesa, a duração de VOT de 23,2ms para /p/ de bilíngues intermediários em inglês (11,62 de desvio-padrão) e 17,96ms para bilíngues avançados em inglês (7,22 de desvio-padrão). No trabalho atual, a média de /p/ em inglês pelas brasileiras bilíngues foi muito maior, 47,6ms. Mais adiante, tem-se que as diferenças de VOT entre brasileiras e estadunidenses são significativas. Porém, como observado na literatura, algumas produções de brasileiras se encaixariam dentro da categoria de VOT observado para nativas estadunidenses, mesmo que na fronteira inferior da duração (próxima de 50ms).

Os valores de VOT de /p/ encontrados na presente pesquisa em português estiveram entre 7ms e 74ms para estadunidenses e 6ms e 51ms para brasileiras (sendo estes valores mais próximos aos de Cohen, 2004). Kupske (2016) encontrou, para monolíngues de português, 15,13ms para /p/, o que é, praticamente, o mesmo valor demonstrado pelas bilíngues brasileiras em português no presente trabalho (14,9ms). Outros valores trazidos por Kupske (2016) para /p/ em português foram: Klein (1999) com 15,59ms; Gewehr-Borella et al. (2011) com 16,6ms; Reis e Nobre-Oliveira (2007) 17,27 ms (desvio-padrão de 6,03); Cohen (2004) encontrou valores entre 14ms e 46ms para /p/ em português. A média de /p/ para monolíngues em Schereschewsky *et al*. (2017) foi de 13,19ms (também bem próximo ao valor de 14,9ms da presente pesquisa).

Os valores de VOT em português produzidos pelas estadunidenses foram superiores às médias da literatura para as monolíngues brasileiras, enquanto as bilíngues brasileiras da presente pesquisa produziram VOTs bem parecidos aos encontrados na literatura. Todavia, essa diferença entre estadunidenses e brasileiras na língua portuguesa não foi significativa pelos testes estatísticos. Guardo a discussão sobre VOT em português e inglês para as próximas seções. Com um olhar pelas lentes da complexidade, retomo Baranger (2000) ao trazer o conceito da não linearidade de um SDC. É possível observar em diversas medidas um grande desvio-padrão (como 22,5 na duração absoluta do VOT de brasileiras em inglês e 15,4 na duração absoluta do VOT de estadunidenses em português) que indicam essa não linearidade.

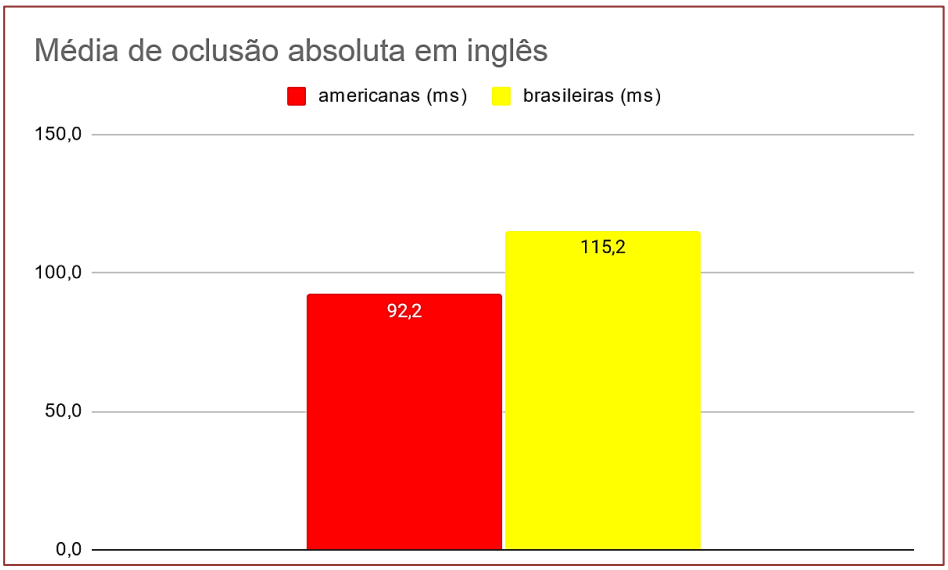
Nas próximas seções, as médias de oclusão, VOT e da duração total da consoante serão observadas com mais detalhes e com uso de testes estatísticos.

6.1.1. MÉDIAS DE DURAÇÃO DE OCLUSÃO

Apresento, primeiramente, as médias e comparações da duração do período de oclusão. Nota-se que as escalas foram pensadas de modo a ficarem as mesmas para cada valor medido de oclusão, VOT e consoante. Por exemplo: o eixo y (vertical, ou das ordenadas) para oclusão em todos os gráficos de duração absoluta da oclusão por meio de gráficos de barra vai até 150ms. As cores adotadas nos gráficos para cada nacionalidade e língua são: vermelho (estadunidenses em inglês), azul (estadunidenses em português), amarelo (brasileiras em inglês) e verde (brasileiras em português). Com o intuito de observar a variação de dados e valores discrepantes, adiciono também gráficos de *boxplot* (diagrama de caixa) para as medidas reportadas. Após as Figuras 6 e 7, ofereço um breve subsídio para a interpretação dos gráficos fornecidos neste trabalho.

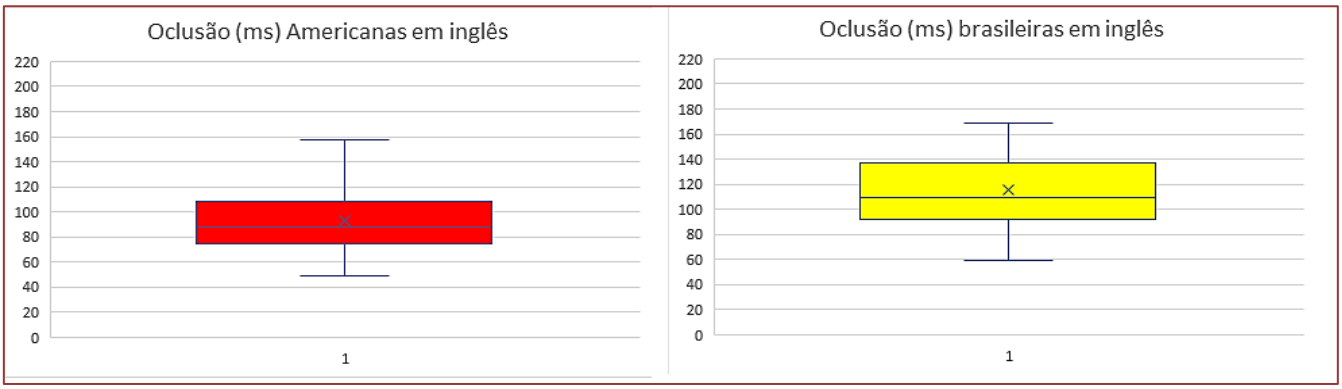
Apresento, nas Figuras 6 e 7, as médias totais da duração de oclusão em *pie* nos textos e sentenças-veículo em inglês (ms).

FIGURA 6 – DURAÇÃO ABSOLUTA DA OCLUSÃO EM INGLÊS POR BRASILEIRAS E ESTADUNIDENSES VIA GRÁFICO DE BARRAS



FONTE: O autor (2021).

FIGURA 7 – DURAÇÃO ABSOLUTA DA OCLUSÃO EM INGLÊS POR BRASILEIRAS E ESTADUNIDENSES VIA DIAGRAMA DE CAIXAS



FONTE: O autor (2021).

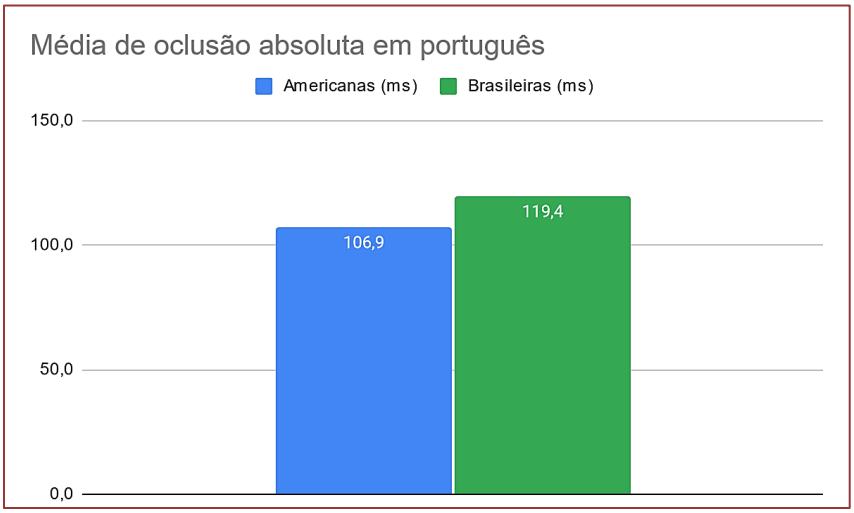
Cada barra na Figura 6 representa a média dos valores de duração absoluta da oclusão (a barra amarela das brasileiras está maior na vertical, pois os valores de duração foram maiores que os valores de duração das estadunidenses). A Figura 7, que apresenta os diagramas de caixa, possui uma linha horizontal dentro de cada caixa que informa a mediana dos valores, ou seja, o valor central de oclusão do conjunto de dados analisados. Quaisquer valores além das outras barras horizontais que se encontram fora das barras (quartil 1 embaixo e quartil 3 em cima) serão descritos como valores discrepantes e podem oferecer pistas interessantes sobre produções desviantes da média de certas informantes. No caso da Figura 7, não existem tais valores discrepantes.

O valor encontrado para duração absoluta da oclusão para as estadunidenses foi de 92,2ms, DP=6ms, enquanto o das brasileiras foi de 115,2ms, DP=12,1ms. Pela Figura 7, tem-se que a mediana da duração absoluta para as estadunidenses foi de 92ms e, para as brasileiras, 112ms. Os valores de duração da oclusão e de VOT são inversamente proporcionais (brasileiras apresentaram VOTs menores e oclusões maiores e vice-versa para as estadunidenses). Estatisticamente, foram encontradas diferenças significativas entre estadunidenses e brasileiras: o teste *Mann-Whitney* apresentou U=0; p<0,001. A média para as estadunidenses foi de 23,3% da palavra, DP=1,95%, enquanto para as brasileiras 28,0%, DP=3,24%. Diferenças significativas foram encontradas mais uma vez: *Mann-Whitney* resultou em U=6; p=0,017. Ademais, esses testes estatísticos estão relacionados à primeira hipótese, que previa que estadunidenses produziram oclusões significativamente mais curtas que as oclusões das informantes brasileiras.

Enfocando especificamente os valores de duração relativa da oclusão, temos que nenhuma falante americana teve valores superiores à média brasileira, bem como nenhuma brasileira teve valores inferiores à média americana. Muitos valores, porém, ficaram próximos à média de sua respectiva nacionalidade. Uma possível contribuição à área da fonética forense para os dados em português é referente ao desvio-padrão das produções. Reforçando a possibilidade de coletar dados em dois idiomas para falantes bilíngues, está o achado de que as informantes apresentaram desvio-padrão muito mais altos em suas respectivas línguas estrangeiras (Estadunidenses em inglês: 6,92ms; português 15,4ms; brasileiras em inglês 22,5ms; português 11,4ms).

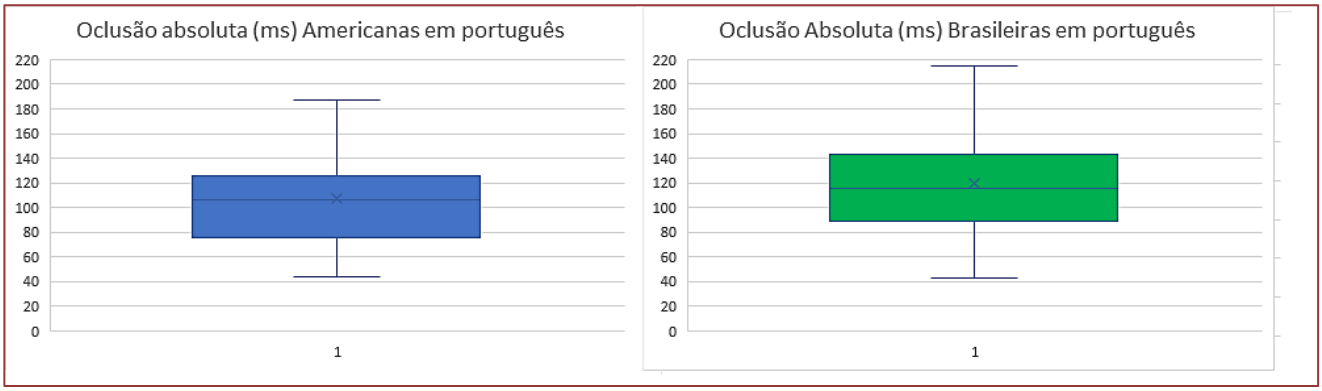
Apresento, nas Figuras 8 e 9, a média da duração da absoluta da oclusão em /p/ na palavra *pai* nas tarefas de leitura de texto e sentenças-veículo em português (ms).

FIGURA 8 – DURAÇÃO ABSOLUTA DA OCLUSÃO EM PORTUGUÊS POR BRASILEIRAS E ESTADUNIDENSES VIA GRÁFICO DE BARRAS



FONTE: O autor (2021).

FIGURA 9 – DURAÇÃO ABSOLUTA DA OCLUSÃO EM PORTUGUÊS POR BRASILEIRAS E ESTADUNIDENSES VIA DIAGRAMA DE CAIXAS



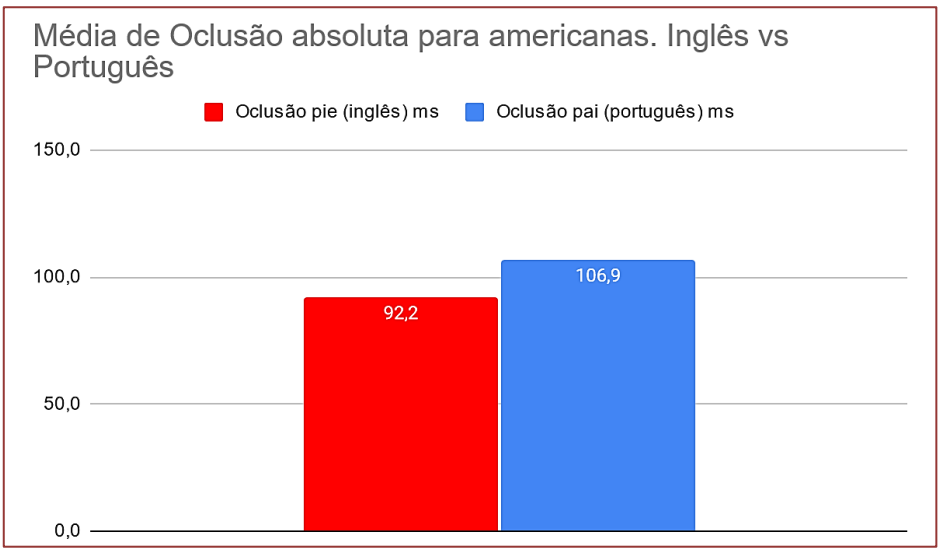
FONTE: O autor (2021).

O valor de duração absoluta de oclusão para as estadunidenses em português foi de 106,9ms, DP=18,9ms, enquanto para as brasileiras o valor encontrado foi 119,4ms, DP=21ms. A mediana das estadunidenses foi de 107ms, e a das brasileiras de 114ms, como é possível observar na Figura 9. Como era esperado, a duração da oclusão das brasileiras em português foi maior do que das estadunidenses (os valores médios encontrados em inglês foram, respectivamente, 92,2ms e 115,2ms). Na língua portuguesa não foram encontradas diferenças significativas pelo teste estatístico[22]. Em relação a média da duração relativa para a oclusão, as estadunidenses apresentaram valores de 26,3% da palavra, DP=9,44%, enquanto o valor para as brasileiras foi 32,8%, DP=3,73%, mantendo a tendência de oclusões maiores para brasileiras. A medida de duração relativa da oclusão em português, contudo, contrariando os testes anteriores do idioma, apresentaram diferenças significativas: *Mann-Whitney U*=3; *p*=0,004. Esses testes estatísticos estão ligados à hipótese 1, que previa oclusões mais longas por parte das brasileiras.

Nas Figuras 10 e 11 seguem as comparações da duração absoluta da oclusão de /p/ para estadunidenses em inglês e português.

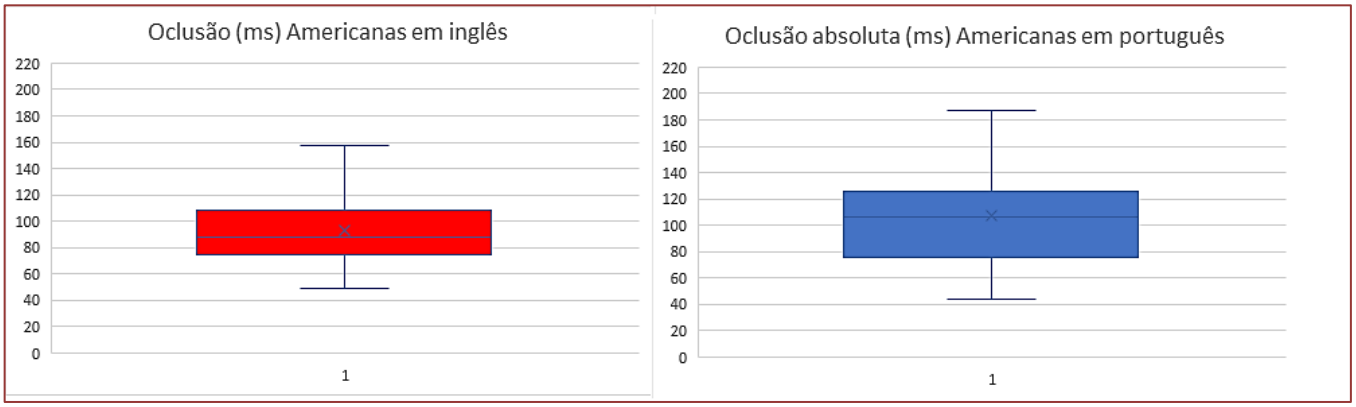
[22]Para manter o texto mais limpo, optei por não reportar os valores não significativos no texto. Todos os resultados dos testes podem ser encontrados nos Apêndices 7 e 8.

FIGURA 10 – DURAÇÃO ABSOLUTA DA OCLUSÃO EM INGLÊS E PORTUGUÊS PARA ESTADUNIDENSES VIA GRÁFICO DE BARRAS



FONTE: O autor (2021).

FIGURA 11 – DURAÇÃO ABSOLUTA DA OCLUSÃO EM INGLÊS E PORTUGUÊS PARA ESTADUNIDENSES VIA DIAGRAMA DE CAIXAS

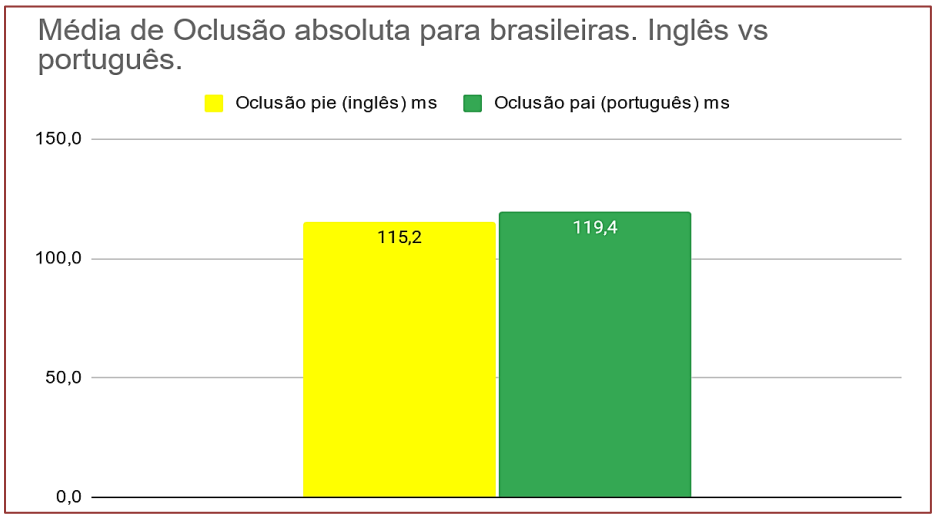


FONTE: O autor (2021).

O valor de duração absoluta da oclusão das estadunidenses em inglês teve a média de 92,2ms, DP=6ms e 106,9ms, DP=18,9ms, em português. O teste de *Wilcoxon* indicou significância com U=28 e p=0,016. O valor de duração relativa da oclusão para as estadunidenses em inglês foi de 23,3%, DP=1,95%, enquanto em português o valor foi de 26,3%, DP=9,44%. Desta vez, o teste estatístico de *Wilcoxon* não indicou significância. Esses testes estão relacionados às hipóteses 2 e 3, que previa que ambas as nacionalidades produziram oclusões mais curtas em inglês do que em português.

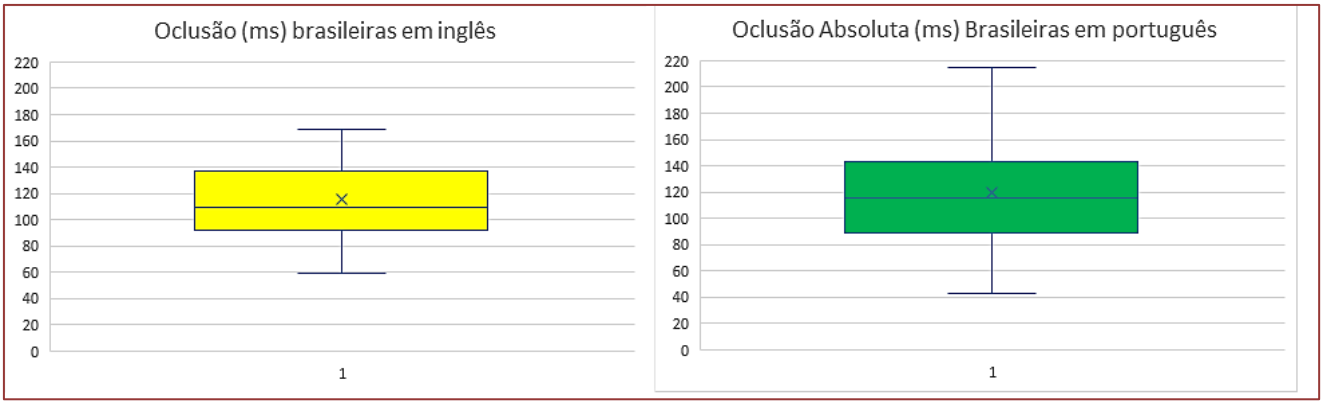
Nas Figuras 12 e 13, apresento a comparação da duração absoluta da oclusão de /p/ para brasileiras em inglês e português.

FIGURA 12 – DURAÇÃO ABSOLUTA DA OCLUSÃO EM INGLÊS E PORTUGUÊS PARA BRASILEIRAS VIA GRÁFICO DE BARRAS



FONTE: O autor (2021).

FIGURA 13 – DURAÇÃO ABSOLUTA DA OCLUSÃO EM INGLÊS E PORTUGUÊS PARA BRASILEIRAS VIA DIAGRAMA DE CAIXAS



FONTE: O autor (2021).

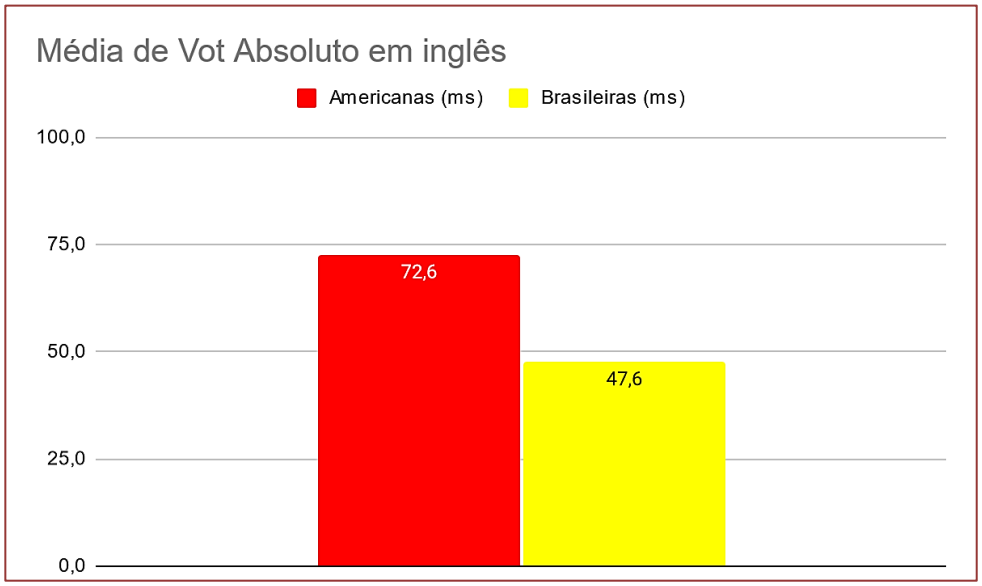
O valor da duração absoluta de oclusão para brasileiras em inglês foi de 115,2ms, DP=12,1ms enquanto em português foi de 119,4ms, DP=21ms. O teste *Wilcoxon* não apresentou diferenças significativas. Já na duração relativa para oclusão, as brasileiras apresentaram 28% da palavra em inglês, DP=3,24% e 32,8% em português, DP=3,73%. Diferentemente dos valores de duração absolutos, aqui diferenças significativas puderam ser observadas pelo teste de *Wilcoxon*, com U=28; p=0,022. Os testes apresentados acima estão relacionados às hipóteses 2 e 3, que previa menores valores de duração da oclusão em inglês para ambas as nacionalidades.

Apresento, em sequência, estatísticas descritivas e inferenciais para analisar os valores de duração de VOT encontrados em inglês e português pelas informantes estadunidenses e brasileiras.

6.1.2. MÉDIAS DE DURAÇÃO DE VOT

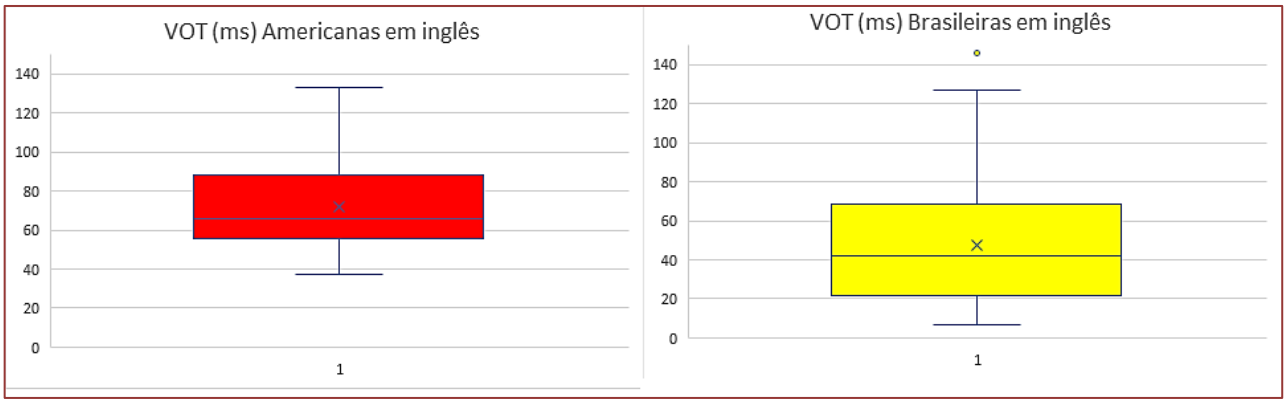
Começo a seção de VOT por comparar estadunidenses e brasileiras em inglês. Nas Figuras 14 e 15 estão a média da duração absoluta do VOT de /p/ em *pie* em todas as ocorrências de textos e sentenças-veículo (ms).

FIGURA 14 – DURAÇÃO ABSOLUTA DO VOT EM INGLÊS POR BRASILEIRAS E ESTADUNIDENSES VIA GRÁFICO DE BARRAS



FONTE: O autor (2021).

FIGURA 15 – DURAÇÃO ABSOLUTA DO VOT EM INGLÊS POR BRASILEIRAS E ESTADUNIDENSES VIA DIAGRAMA DE CAIXAS



FONTE: O autor (2021).

O valor total médio de VOT em *pie* foi de 72,6ms, DP=6,92ms, para estadunidenses e 47,6ms, DP=22,5ms, para brasileiras. Isso indica a tendência de as estadunidenses produzirem VOTs mais longos. A mediana de VOT das estadunidenses ficou em 72ms, enquanto para as brasileiras se tem 49ms. A Figura 15 apresenta uma interessante

produção discrepante por parte de uma informante do grupo de brasileiras: houve uma ocorrência acima dos 140ms, duração essa que não ocorreu nem mesmo pelas informantes estadunidenses (que registraram VOTs de durações mais longas por média e mediana). A partir de alguns testes estatísticos, é possível encontrar diferenças significativas entre as produções de brasileiras e estadunidenses, pois o teste *Mann-Whitney* indicou U=8; p=0,041. O valor de duração relativa média de VOT para estadunidenses foi de 18,1% da palavra, DP=1,46%, enquanto para as brasileiras o valor encontrado foi 11,5%, DP=4,82%. A tendência de VOT maior para estadunidenses se manteve também nessa medida. Assim como nas durações absolutas em milissegundos, os valores de duração relativa de VOT apresentaram diferenças: para o teste de *Mann-Whitney*, U=1; p=0,03. Esses primeiros testes estatísticos colaboram para a verificação da hipótese 1, no sentido de clarificar a diferença significativa de VOT entre brasileiras e estadunidenses em inglês (sendo mais longas as produções de estadunidenses em inglês). Portanto, é possível afirmar que os grupos de nacionalidade apresentam diferenças significativas nas durações absolutas e relativas de VOT. É importante relembrar que todas as 14 informantes declararam fazer uso das duas línguas em seu cotidiano e serem fluentes em português e inglês. As brasileiras deste estudo, não obstante, tendem a produzir valores de VOT significativamente mais baixos do que o padrão nativo. Seria interessante que fossem feitas pesquisas no campo de percepção para averiguar se essa diferença nos valores de VOT realmente influenciam o ouvinte e a inteligibilidade. Caso haja problemas de inteligibilidade e percepção, a questão da produção do VOT precisaria, talvez, ser mais trabalhada em sala de aula. Caso contrário, temos apenas uma possível característica da produção de brasileiras falando inglês, o que pode ser de grande utilidade para a área de comparação de locutores dentro da fonética forense, uma vez que características que individualizam um falante são essenciais para determinações legais e de segurança. Um parâmetro como o do VOT pode ser útil para selecionar um grupo menor de possíveis identidades dentro de um amplo universo de informantes. A possível contribuição forense dessa parte do trabalho está em como as durações de VOT podem individualizar cada informante. Somente uma informante brasileira (B18) manteve 79,6ms, uma média de duração absoluta do VOT em /p/ maior que a média das estadunidenses como grupo de nacionalidade, a cargo de exemplo. Vale lembrar, conforme comentado na Seção 6.1, que muitos valores das brasileiras falando em inglês se encaixariam nas durações de VOT encontradas para nativos do inglês, e essa variabilidade poderá ser observada nos gráficos de dispersão, que são encontrados mais adiante no texto. De todo modo, essa diferença passa a ser interessante em termos de comparação de locutores.

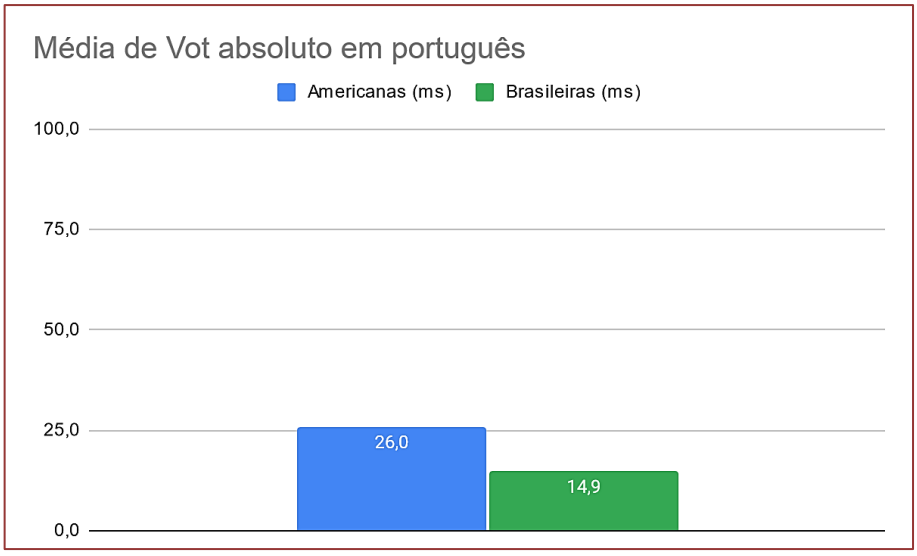
Outro detalhe a ser levado em conta é que as informantes brasileiras desta pesquisa residem no Brasil, enquanto as brasileiras da pesquisa de Kupske (2016) residiam na Inglaterra. As estadunidenses da presente pesquisa *também* moravam no Brasil, então poderia ser interessante fazer uma outra comparação com brasileiras morando na costa leste dos Estados Unidos (região da qual a maioria das informantes estadunidenses era proveniente). No caso da pesquisa de Kupske (2016), comprovou-se que as línguas faladas pelas brasileiras que moravam em Londres acabavam se alterando (por meio do processo de atrito linguístico) com o passar do tempo de residência no país europeu. Talvez, o que observamos nos dados do trabalho atual seja o processo de atrito em maior ou menor grau por parte das informantes estadunidenses.

O quinto tópico levantado por Baranger (2000) trata da relação de caos com não caos existente em um SDC. É observável a existência de um parâmetro de controle que

evita o transbordamento do complexo para o caótico: nenhuma informante produziu VOTs, por exemplo, de 1ms ou 150ms. Elas sempre transitaram entre determinadas fronteiras de valores e, geralmente, seus valores se agrupavam em uma faixa restrita de valores. Apenas uma informante dentre as 14 (identificada como B18, no texto em inglês) apresentou durações absolutas de VOTs em todos os intervalos de duração diferentes[23], apresentando valores de 19ms, 41ms, 53ms, 77ms, 94ms, 127ms e 146ms. A regra geral para as informantes foi apresentar VOTs em dois ou três intervalos de duração. A informante A4, inclusive, apresentou todas as produções de duração absoluta de VOT na mesma faixa de valores (62ms, 63ms, 65ms, 74ms, 77ms, 83ms e 86ms), apresentando um DP muito pequeno. Inclusive, sete das 14 informantes produziram VOTs na mesma faixa de duração em português (A5, A8, B14, B15, B17, B18 e B20). Fica clara a tendência de produções escaparem, mas não serem totalmente caóticas.

Encontram-se, nas Figuras 16 e 17, as médias totais do VOT de /p/ em *pai* nos textos e sentenças-veículo em português (ms).

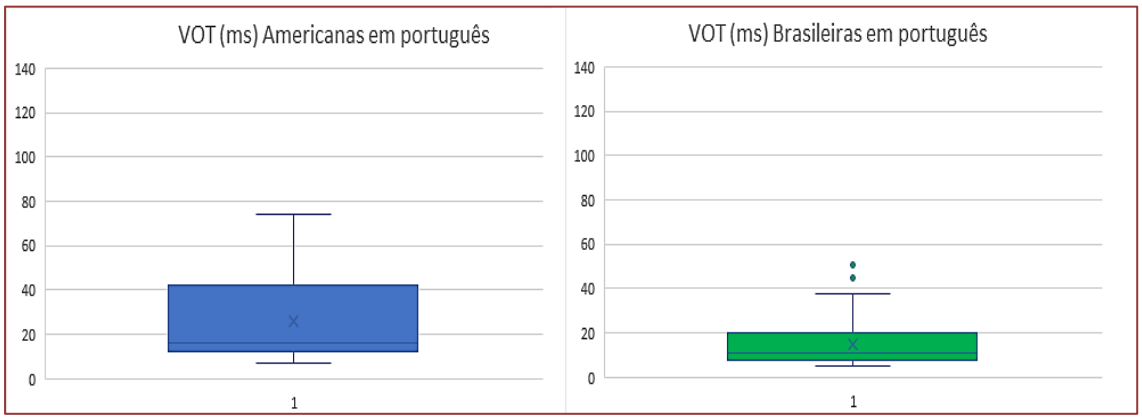
FIGURA 16 – DURAÇÃO ABSOLUTA DO VOT EM PORTUGUÊS POR BRASILEIRAS E ESTADUNIDENSES VIA GRÁFICO DE BARRAS



FONTE: O autor (2021).

[23]Com finalidades organizacionais, adotei em minhas planilhas de dados diferentes categorias de duração de VOT: VOTs de duração curta de até 20ms, VOTs médios de 20ms até 60ms, VOTs médio-longo de 60ms até 100ms e VOTs longos maiores que 100ms.

FIGURA 17 – DURAÇÃO ABSOLUTA DO VOT EM PORTUGUÊS POR BRASILEIRAS E ESTADUNIDENSES VIA DIAGRAMA DE CAIXAS



FONTE: O autor (2021).

O valor médio de VOT, em português, para *pai* das estadunidenses foi de 26ms, DP=15,4ms, enquanto para as brasileiras temos 14,9ms, DP=11,4ms. Pela Figura 17, é possível perceber que a mediana das estadunidenses ficou em 22ms, enquanto a das brasileiras ficou em 12ms. A Figura 17 também dispõe duas produções discrepantes que representam VOTs superiores a 40ms no português por informantes do grupo de brasileiras. Esse fato é interessante, pois é um valor muito alto para nativas de português brasileiro e fica bem acima até mesmo da média de estadunidenses. É possível observar que os valores de duração absoluta de VOT são muito menores que os 72,6ms e 47,6ms do inglês, porém a tendência de um VOT mais longo por parte das estadunidenses se mantém. Diferentemente do inglês, porém, em que o teste estatístico apontou diferenças significativas entre estadunidenses e brasileiras, o VOT em português não apresentou os mesmos resultados. Esses resultados também implicam na verificação da hipótese 1, relacionada, entre outros fatores, nas diferenças entre nacionalidades nas produções de VOT. Também chamo atenção ao fato de os DPs para cada nacionalidade serem maiores na língua estrangeira: estadunidenses em inglês: 6,92ms, em português 15,4ms; brasileiras em inglês 22,5ms, em português 11,4ms. O valor da duração relativa do VOT para as estadunidenses foi de 6,3%, DP=2,29%, enquanto o das brasileiras foi de 4,4%, DP=3,02%. Nota-se que os valores das estadunidenses continuam mais altos que o das brasileiras mesmo na média da duração relativa. As durações relativas de VOTs em português, assim como as absolutas, não apresentaram diferenças significativas em nenhum teste estatístico analisado. Os resultados foram bem diferentes dos resultados observados em língua inglesa. Isso quer dizer que, utilizando-se dos parâmetros de duração absoluta de VOT, as brasileiras e estadunidenses estariam produzindo, virtualmente, a mesma consoante em português, mas não em inglês. Talvez, isso se dê pelo fato de as estadunidenses residirem no Brasil, mas as brasileiras não residirem em um país de língua inglesa no momento da coleta dos dados, como mencionado no parágrafo discutindo o atrito na pesquisa de Kupske (2016).

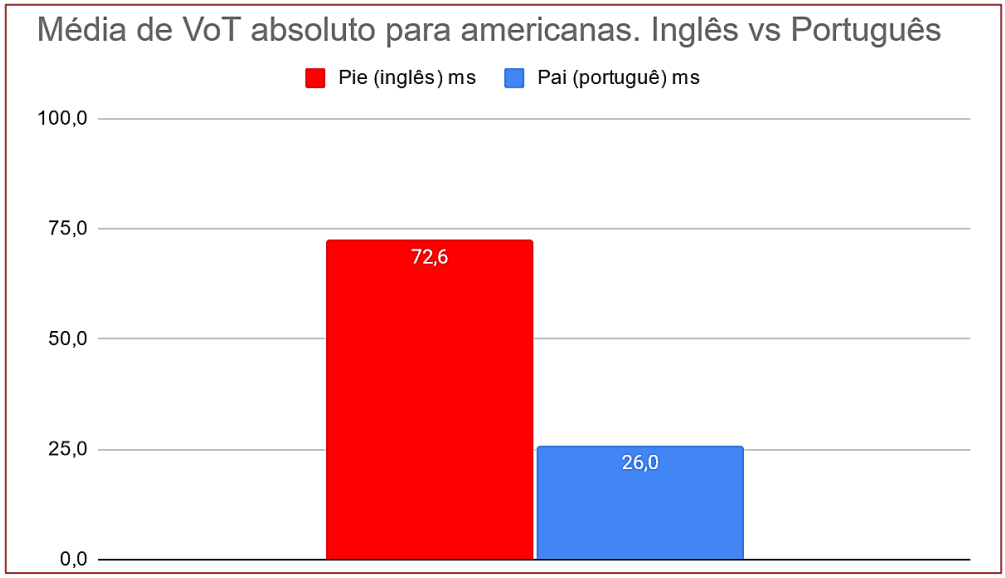
Retomando a ideia de complexidade, mais especificamente no segundo tópico apresentado por Baranger (2000), que trata da interdependência, nota-se que essa característica pode ser vista por meio das diferentes médias de cada informante. O mesmo pode ser exemplificado no português com as informantes B7 (10ms de média em

português), B12 (27,3ms) e B17 (6,8ms). Cada informante é um agente interdependente no sistema linguístico de cada língua. A sexta relação observada por Baranger (2000), sobre a relação entre o competitivo e o cooperativo, fica clara nos desvio-padrão de certas informantes. A informante A6 produziu VOTs médios e médio-longos em inglês (50ms, 51ms, 57ms, 76ms, 88ms, 89ms, 98ms) e curtos e médios em português (16ms, 16ms, 20ms, 21ms, 21ms, 42ms), estabelecendo categorias de valores que, apesar de definidas, competem entre si. O quarto tópico trazido por Rzevski (2012), sexto trazido por Chan (2001), se refere aos eventos disruptivos que impedem um SDC de estar sempre em equilíbrio. Talvez, a instrução explícita de VOT em sala de aula seja um evento disruptivo para um aprendiz de inglês como L2 ou a súbita percepção de um estadunidense que as durações de VOT em português tendem a ser menores.

Apresento, na sequência, os resultados de testes pareados que averiguam as nacionalidades individualmente. Comparo informantes de uma mesma nacionalidade produzindo segmentos das duas línguas. Os testes estatísticos desta seção foram pareados não paramétricos (*Wilcoxon* e não *Mann-Whitney*) e a significância continua sendo restrita a valores de *p* menores ou iguais a 0,05.

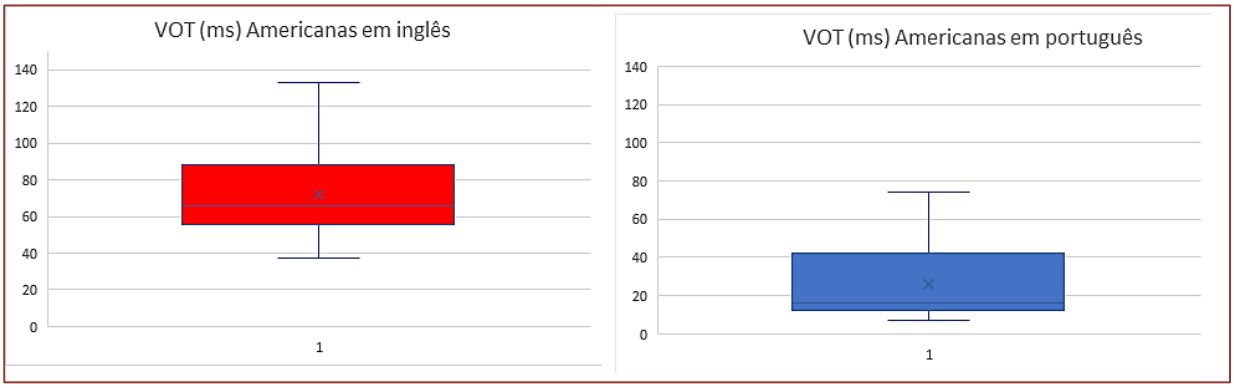
Nas Figuras 18 e 19 estão a comparação de duração absoluta do VOT de /p/ para estadunidenses em inglês e português.

FIGURA 18 – DURAÇÃO ABSOLUTA DO VOT POR ESTADUNIDENSES EM PORTUGUÊS E INGLÊS VIA GRÁFICO DE BARRAS



FONTE: O autor (2021).

FIGURA 19 – DURAÇÃO ABSOLUTA DO VOT POR ESTADUNIDENSES EM PORTUGUÊS E INGLÊS VIA DIAGRAMA DE CAIXAS



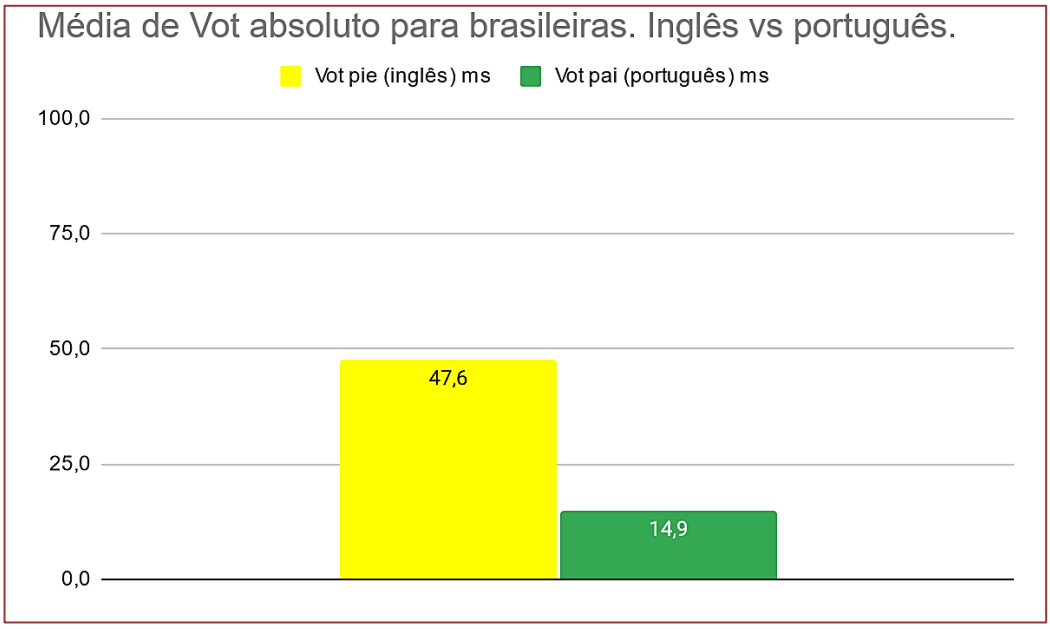
FONTE: O autor (2021).

A duração absoluta média do VOT das estadunidenses em inglês foi de 72,6ms, DP=6,92ms, enquanto em português o valor foi de 26ms, DP=15,4ms. O teste estatístico *Wilcoxon* (U=0; p=0,016) apontou diferenças significativas entre os valores. A duração relativa do VOT das estadunidenses em inglês foi de 18,1 (porcentagem na duração da palavra), DP=1,36%, enquanto em português o valor foi 6,2%, DP=2,29%. Para essa medida, o teste de *Wilcoxon* também apontou diferenças significativas, U=0; p=0,016. Esses resultados ajudam a elucidar as hipóteses 2 e 3, que discutem, entre outras especificidades, se cada nacionalidade produz os VOTs em inglês de forma mais longa que os VOTs em português.

Até aqui, nota-se que as diferenças significativas foram encontradas no VOT entre as línguas. As estadunidenses confirmaram a tendência de diferenciar os VOTs entre português e inglês, mas essa padronização não se estendeu na mesma proporção para os períodos de oclusão (exceto em um teste de duração absoluta da oclusão). Se as medidas de duração absoluta de VOT e duração absoluta da oclusão entre as línguas diferem tanto, ambas podem ser eventuais pistas de individualização.

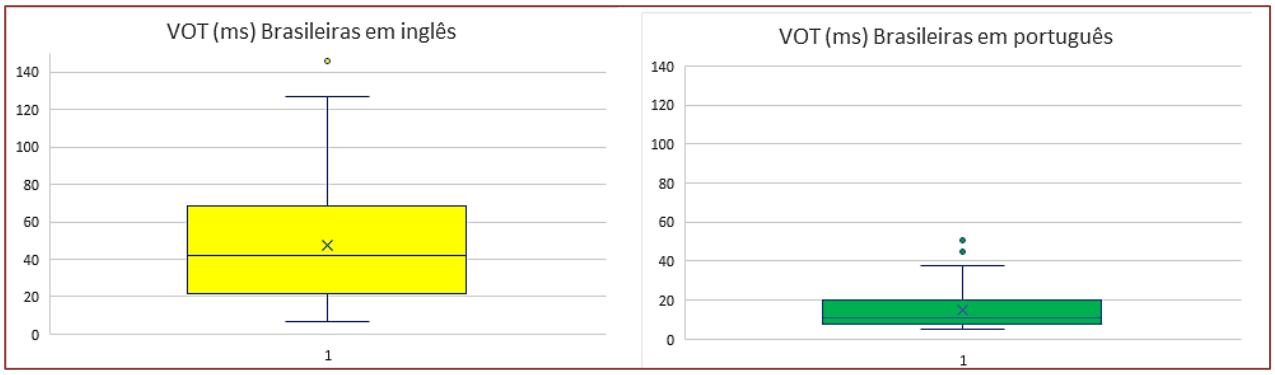
Encontram-se, nas Figuras 20 e 21, as comparações de duração absoluta de VOT de /p/ para brasileiras em inglês e português.

FIGURA 20 – DURAÇÃO ABSOLUTA DO VOT POR BRASILEIRAS EM PORTUGUÊS E INGLÊS VIA GRÁFICO DE BARRAS



FONTE: O autor (2021).

FIGURA 21 – DURAÇÃO ABSOLUTA DO VOT POR BRASILEIRAS EM PORTUGUÊS E INGLÊS VIA DIAGRAMA DE CAIXAS



FONTE: O autor (2021).

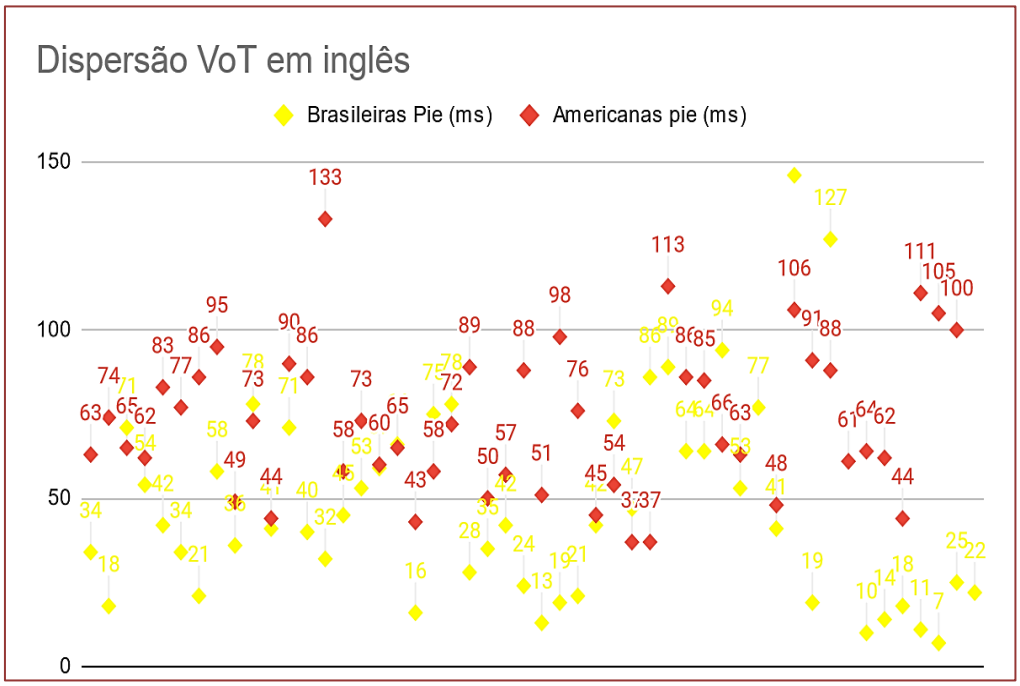
O valor médio de duração absoluta do VOT para as brasileiras em inglês foi de 47,6ms, DP=22,5ms, enquanto em português encontrei 14,9ms, DP=11,4ms. Nota-se, pela Figura 21, que em inglês houve uma produção superior a 140ms (fato já comentado na análise entre nacionalidades). Em português, notam-se as duas produções discrepantes acima dos 40ms. É interessante perceber que o grupo de brasileiras aparentou ter produções de durações muito mais heterogêneas do que o grupo de estadunidenses. O teste estatístico *Wilcoxon* indicou resultados significativos, sendo U=2; p=0,047. A duração relativa do VOT para brasileiras em inglês foi de 11,5 (porcentagem da palavra), DP= 4,82%, enquanto em português foi 4,4%, DP=3,02%. O teste de *Wilcoxon* (U=0; p=0,016) também indicou diferenças significativas. Esses testes também colaboram para

a hipótese 3, uma vez que medem a significância de valores das brasileiras entre as duas línguas.

Como observado, as brasileiras apresentaram diferenças significativas entre a duração absoluta/relativa de VOT no inglês e no português. É interessante observar o fato que houve diferenças significativas entre os grupos de nacionalidade e houve diferença intragrupos comparando as duas línguas. Nota-se também que, embora as brasileiras produzam um VOT mais curto do que as estadunidenses em inglês, o VOT das brasileiras em inglês e português são diferentes (se os valores de oclusão e VOT são diferentes, não é exatamente a mesma consoante /p/ nas duas línguas).

Para melhor visualização das diferentes produções dos grupos de nacionalidade, disponho gráficos contendo a dispersão de duração absoluta dos valores de VOT, primeiramente em inglês, na Figura 22.

FIGURA 22 – DISPERSÃO DA DURAÇÃO ABSOLUTA DO VOT EM INGLÊS POR AMBAS NACIONALIDADES

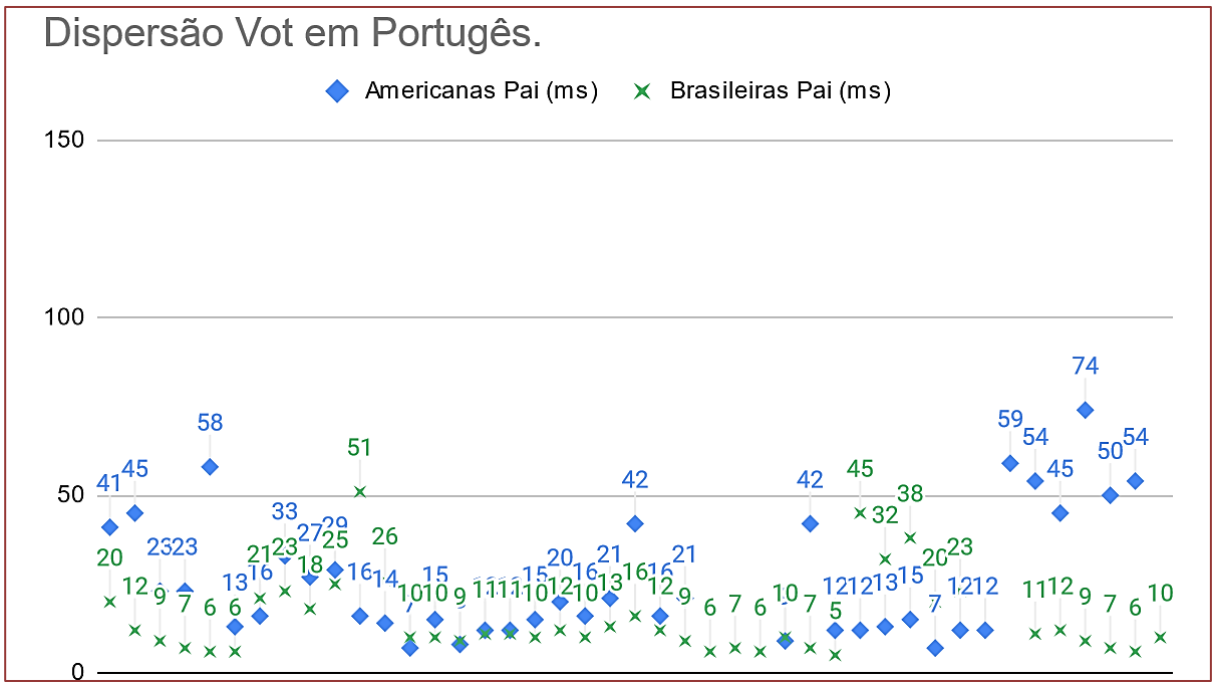


FONTE: O autor (2021).

Ao analisar a dispersão de todas as produções quantificadas de VOT em inglês, nota-se a predominância dos dados das brasileiras nos valores inferiores a 50ms. Enquanto só são observadas sete produções de estadunidenses abaixo de 50ms, as brasileiras contribuem com 30 produções. Entre 50ms e 100ms, as produções de estadunidenses predominam: 34 contra 17 de brasileiras. Na faixa acima dos 100ms, as estadunidenses aparecem com seis resultados, enquanto as brasileiras com apenas dois. As produções mais altas de VOT, todavia, pertencem às brasileiras. A faixa entre 50ms e 100ms é a mais equilibrada, enquanto a porção com VOTs mais curtos é composta, majoritariamente, por brasileiras, a porção com VOTs mais longos é composta, majoritariamente, por estadunidenses.

Apresento, por meio da Figura 23, a dispersão de duração absoluta de /p/ em português (*pai*) por ambas as nacionalidades.

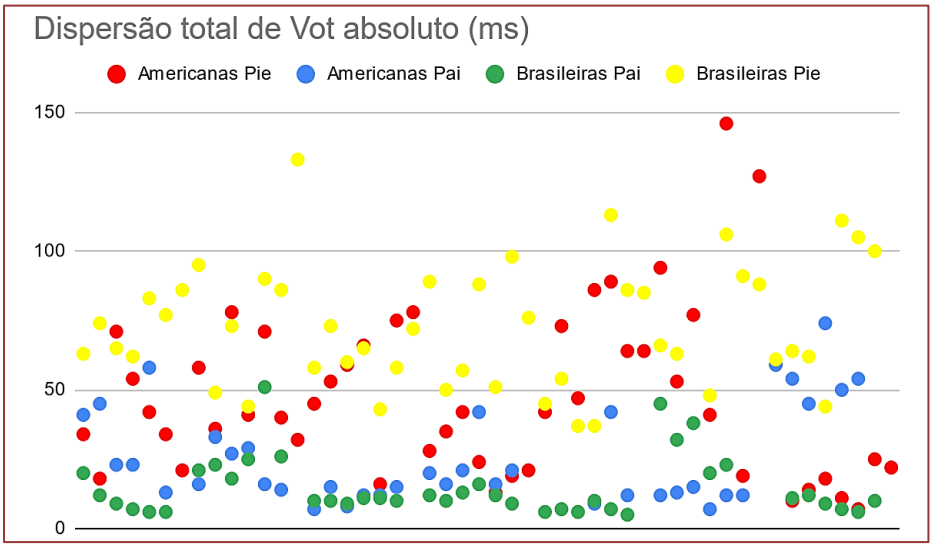
FIGURA 23 – DISPERSÃO DO VOT EM PORTUGUÊS POR AMBAS NACIONALIDADES



FONTE: O autor (2021).

Em português, a faixa entre 0ms e 50ms é composta por 32 produções de estadunidenses e 40 produções de brasileiras. Acima de 50ms, apenas uma produção de brasileira e quatro produções de estadunidenses. É interessante observar que, apesar de haver uma homogeneidade maior do que no gráfico do inglês, a parte de VOTs mais longos continua sendo dominada por estadunidenses. Na Figura 24, se encontra o gráfico de Dispersão total dos valores de VOT.

FIGURA 24 – DISPERSÃO DE DURAÇÃO ABSOLUTA DO VOT NAS DUAS LÍNGUAS/NACIONALIDADES



FONTE: O autor (2021).

Com relação ao gráfico de dispersão total dos VOTs, nota-se a predominância das brasileiras falando português na parte inferior do gráfico (VOTs de durações mais curtas), com muitas produções de estadunidenses também ocupando esse espaço e áreas mais próximas à região de 50ms. A região intermediária (entre 50ms e 100ms) contém a maior parte das produções em inglês, principalmente as realizadas pelas estadunidenses, e a parte superior a 100ms contendo mais estadunidenses que brasileiras, mas com os dois valores mais altos tendo sido produzidos por brasileiras. Assim como em Lima Jr (2015), conforme foi estudado na Seção 4.3, muitas informantes apresentaram produções muito distantes das médias do grupo e das outras próprias produções.

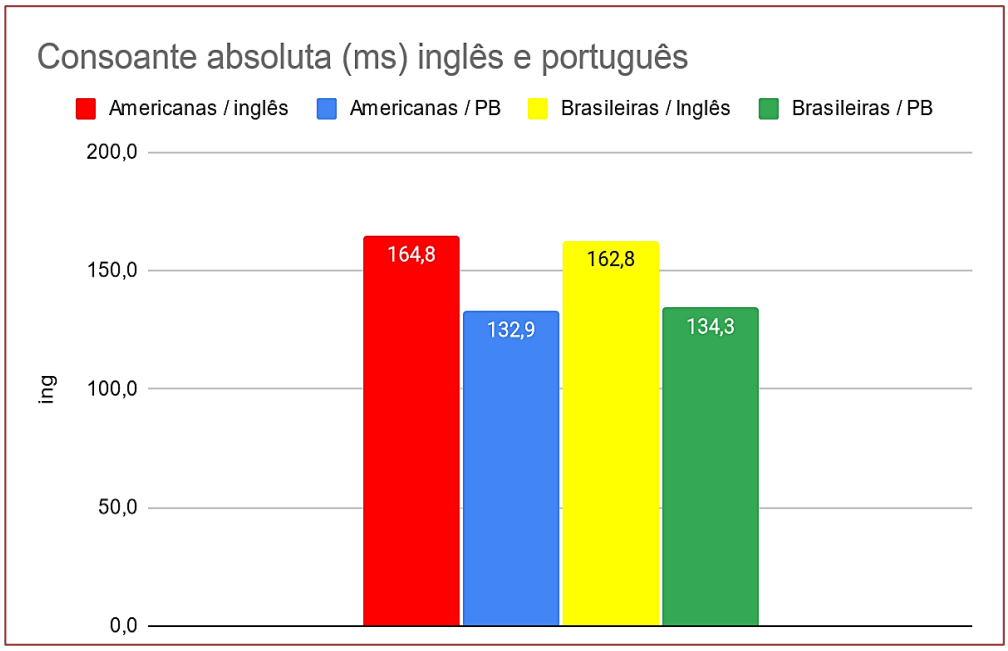
A seguir, apresento os valores de duração total da consoante.

6.1.3. MÉDIAS DE DURAÇÃO TOTAL DE CONSOANTE

Em uma primeira vista, os valores de duração de VOT e oclusão aparentam ser inversamente proporcionais e por isso o valor de duração total da consoante foi calculado (se espera, portanto, que este não seja significativamente diferente entre grupos de nacionalidade e língua, por abarcar ambos VOT e oclusão). A tabela com as médias individuais de duração total da consoante para cada informante se encontra nos Apêndice 5 e 6.

Na Figura 25, demonstro a comparação de duração absoluta total da consoante de /p/ por estadunidenses e brasileiras em inglês e português.

FIGURA 25 – DURAÇÃO ABSOLUTA TOTAL DA CONSOANTE NAS DUAS LÍNGUAS/NACIONALIDADES

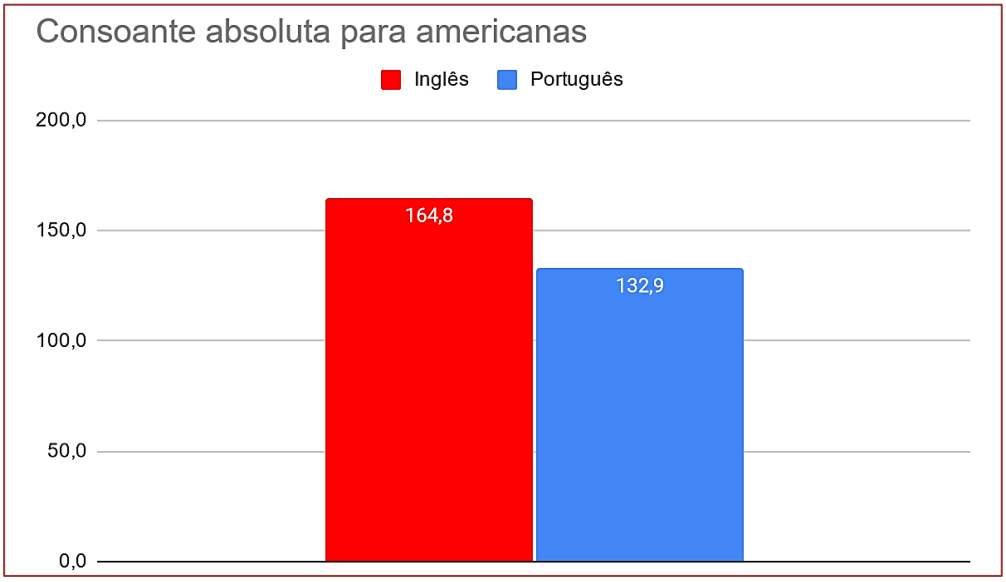


FONTE: O autor (2021).

É possível, por meio da Figura 32, observar o quão próximos são os valores de duração total da consoante nas duas línguas. Mesmo com as diferenças entre duração de VOT e oclusão apresentadas nas primeiras seções deste capítulo, os valores entre estadunidenses e brasileiras em inglês e estadunidenses e brasileiras em português são praticamente os mesmos. Isso será discutido na retomada das hipóteses na Seção 6.3. Não foram encontradas diferenças significativas entre brasileiras e estadunidenses em inglês (duração absoluta e relativa) e em português (absoluta). A única diferença significativa encontrada nos testes foi para estadunidenses/brasileiras em duração relativa total da consoante em português (*Mann-Whitney*, U=7; p=0,026). Dito isso, passo para as análises intranacionalidades de duração total de consoante para cada língua.

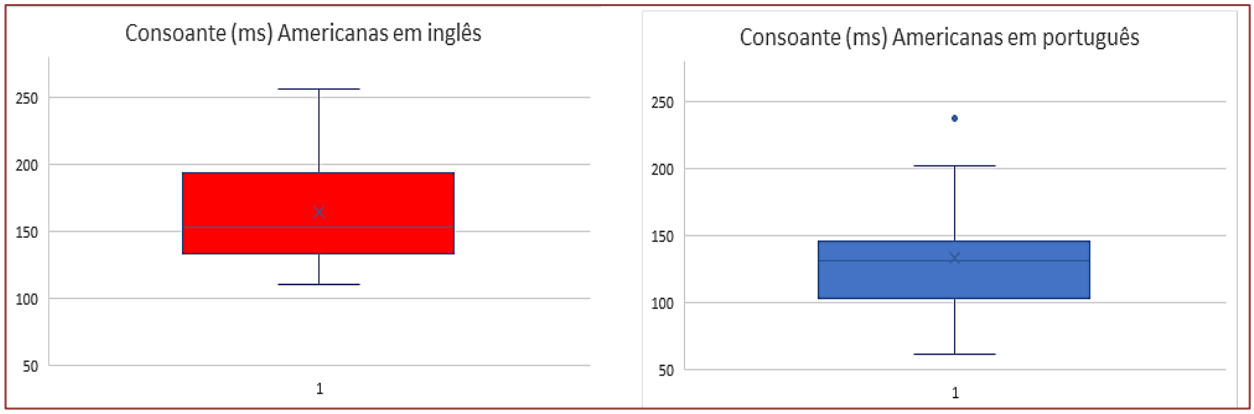
Apresento, por meio das Figuras 26 e 27, a comparação da duração absoluta total da consoante para estadunidenses, inglês *vs* português.

FIGURA 26 – DURAÇÃO ABSOLUTA TOTAL DA CONSOANTE EM AMBAS AS LÍNGUAS PARA ESTADUNIDENSES VIA GRÁFICO DE BARRAS



FONTE: O autor (2021).

FIGURA 27 – DURAÇÃO ABSOLUTA TOTAL DA CONSOANTE EM AMBAS AS LÍNGUAS PARA ESTADUNIDENSES VIA DIAGRAMA DE CAIXAS

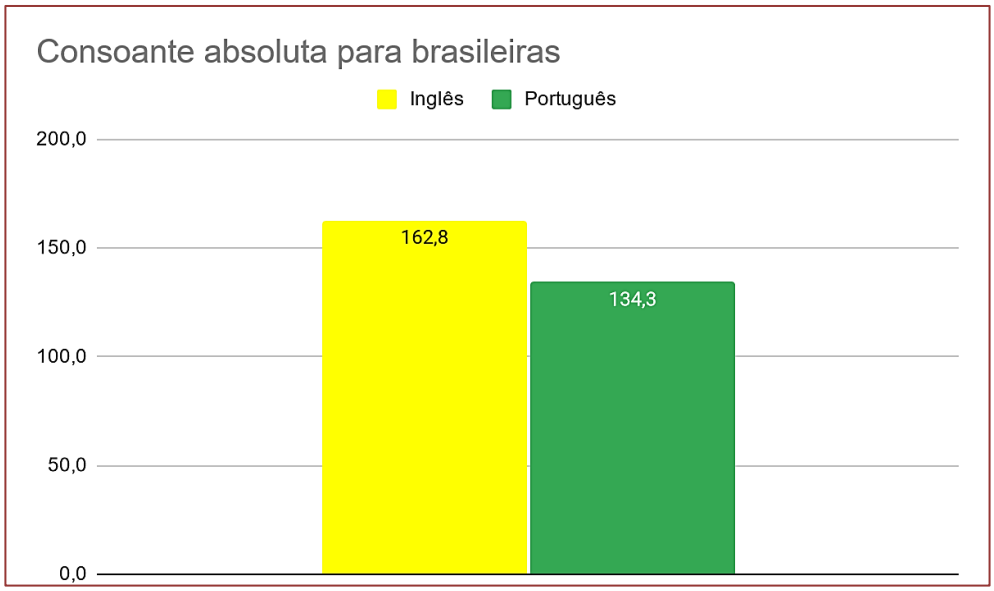


FONTE: O autor (2021).

O valor de duração absoluta total da consoante encontrado para estadunidenses em inglês foi de 164,8ms, DP=9,52ms, enquanto em português o valor foi de 132,9ms, DP=31,2ms. A mediana das estadunidenses em inglês ficou em 160ms, enquanto em português a mediana da duração total da consoante ficou em 133ms. A Figura 27 também mostra uma produção discrepante em português. O teste de *Wilcoxon* não indicou diferenças significativas. O valor de duração relativa para estadunidenses em inglês foi de 41,8 (porcentagem da palavra), DP=2,87%, enquanto em português o valor foi de 32,9%, DP=3,02%. Na duração relativa, o teste de *Wilcoxon* apresentou diferenças significativas, com U=0; p=0,016, o que indica que, pelo menos na duração relativa, as estadunidenses produzem consoantes diferentes nas duas línguas.

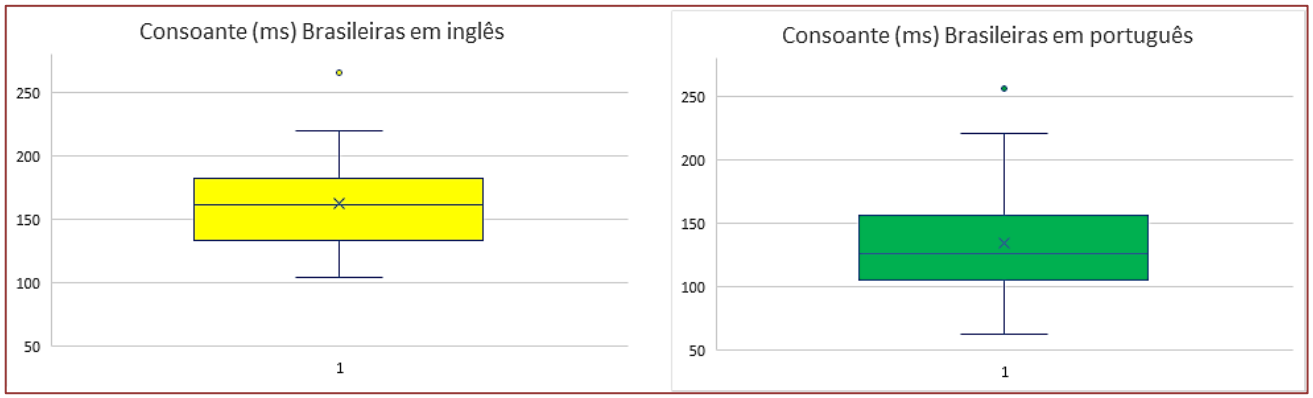
Nas Figuras 28 e 29, está presente a comparação da duração absoluta total da consoante /p/ para brasileiras em inglês e português.

FIGURA 28 – DURAÇÃO ABSOLUTA TOTAL DA CONSOANTE EM AMBAS AS LÍNGUAS PARA BRASILEIRAS VIA GRÁFICO DE BARRAS



FONTE: O autor (2021).

FIGURA 29 – DURAÇÃO TOTAL DE CONSOANTE EM AMBAS AS LÍNGUAS PARA BRASILEIRAS VIA DIAGRAMA DE CAIXAS



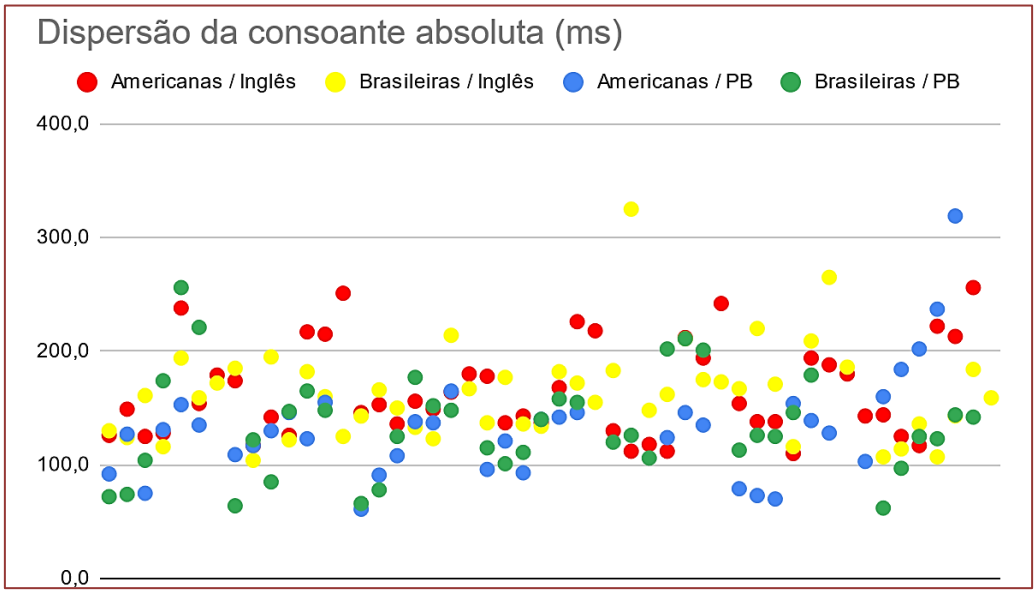
FONTE: O autor (2021).

O valor de duração absoluta total da consoante encontrado pelas brasileiras em inglês foi de 162,8ms, DP=22,3ms, enquanto em português o valor foi de 134,3ms, DP=16,4ms. A mediana das brasileiras em inglês ficou em 156ms e em português em 130ms. Conforme mostra a Figura 29, houve valores discrepantes em ambas as línguas (acima de 250ms em português e inglês). O teste de *Wilcoxon*, com U=0 e p=0,016, indicou significância, ou seja, as brasileiras produziram consoantes de durações diferentes em português e inglês. O valor de duração relativa total da consoante encontrado pelas

brasileiras em inglês foi de 39,5 (porcentagem da palavra), DP=3,13% enquanto em português, DP =2,88%, o valor foi de 37,2%. O teste estatístico para a duração relativa não apontou diferenças significativas, o que mostra que as brasileiras, diferentemente das estadunidenses, não apresentaram diferenças na duração absoluta total da consoante /p/ nas duas línguas.

Apresento os gráficos de dispersão da duração absoluta total das consoantes na Figura 30.

FIGURA 30 – DISPERSÃO DE DURAÇÃO ABSOLUTA TOTAL DA CONSOANTE NAS DUAS LÍNGUAS/NACIONALIDADES



FONTE: O autor (2021).

A Figura 30 representa a grande variabilidade da duração de resultados de duração total da consoante. Nota-se uma grande variedade de produções por ambas nacionalidades e línguas na faixa entre 100ms e 200ms, mais ocorrências de produções de brasileiras com consoantes de durações menores que 100ms e estadunidenses com consoantes maiores que 200ms

Concluindo esta seção, tem-se que as durações totais de consoantes das brasileiras foram significativamente diferentes nas duas línguas apenas ao olhar para durações de valores absolutos (ms), mas não em durações relativas. Para as estadunidenses, os valores de duração relativa total de consoante apresentaram diferenças significativas, mas somente um teste de duração absoluta total da consoante indicou para esse caminho. Todavia, notou-se que os valores quase iguais apresentados na Figura 32 apontam que há alguma evidência para que o VOT e a oclusão sejam grandezas inversamente proporcionais no contexto da pesquisa atual.

6.1.4. COMPARAÇÃO DE TAREFAS: TEXTOS E SENTENÇAS-VEÍCULO

Nesta seção, comparo as diferenças de valores entre os métodos de coleta (texto em português e inglês e sentenças em português e inglês). De acordo com a hipótese 4, que previa diferenças entre as tarefas, pressuponho que as durações absolutas de VOT nas palavras das sentenças-veículo sejam mais longas que as localizadas no texto, uma vez que metade das ocorrências no texto estavam no meio de frases em posições não acentuadas pelos falantes. Por outro lado, espera-se que as durações relativas sejam similares (uma vez que o valor de duração relativa está relacionado à duração total da palavra, tendendo a se normalizar). Não é o objetivo do trabalho realizar uma análise prosódica, então apenas apresento os valores obtidos contidos na Tabela 4. Fora dos parênteses se encontra a duração absoluta (milissegundos) e, entre parênteses, a duração relativa (porcentagem da palavra).

TABELA 4 – DURAÇÕES DE VOT, OCLUSÃO E DURAÇÃO TOTAL DA CONSOANTE ABSOLUTAS E RELATIVAS NOS TEXTOS E SENTENÇAS POR AMBAS AS NACIONALIDADES

		VOT		OCLUSÃO		CONSOANTE	
Part.	Tarefa	ING	PORT	ING	PORT	ING	PORT
AM	SENT	82,7 (18,4)	25,9 (5,5)	110,8(24,9)	131(30,2)	193,4(43,2)	156,9(35,7)
	TEXT	60,5 (17,4)	26,1 (7,1)	75,7 (21,9)	78,8(22,6)	136,2(40,1)	104,9(29,7)
BRAS	SENT	49,2 (11,1)	13,9 (3,4)	127,3(28,4)	155,6(37,1)	176,5(39,5)	169,5(40,6)
	TEXT	46,7 (12,0)	15,9 (5,4)	105,8(27,8)	85(28,6)	152,5(39,8)	100,8(34)

FONTE: O autor (2021).

Para averiguar se as observações da hipótese são confirmadas ou refutadas estatisticamente, fiz o uso do teste estatístico de *Mann-Whitney* para comparar os instrumentos de coleta[24]. Ao se deparar com a Tabela 4, verifica-se que os valores de duração absoluta/relativa de VOT das estadunidenses foram sempre maiores em inglês do que em português e sempre maiores do que os valores das brasileiras. O valor de duração absoluta do VOT das estadunidenses nas sentenças em inglês foi maior do que no texto em inglês. Esse foi o único resultado estatisticamente significativo para VOT entre os instrumentos (U=5,5; p=0,018). As brasileiras também apresentaram essa tendência, apenas de forma menos acentuada e o teste não apontou significância. As durações relativas de VOTs foram maiores também no texto ao se comparar com sentenças (em português) para brasileiras e estadunidenses. Para as brasileiras, os valores de VOT também foram maiores no texto em inglês, mas esta diferença não foi estatisticamente significativa.

Resultados significativos que apontavam diferenças entre as tarefas apareceram para o período de oclusão: as durações absolutas das oclusões das estadunidenses em inglês foram, significativamente, diferentes entre texto e sentenças-veículo (sendo as durações das oclusões das sentenças maiores que as dos textos), com U=0; p=0,02. As durações absolutas das oclusões das brasileiras em inglês também foram significativas, com U=7; p=0,029, o que indica que texto e sentenças-veículo foram diferentes nesse caso também. Em português, apenas as durações relativas de oclusões das estadunidenses não foram significativas. Nessa língua, o resultado de duração absoluta das estadunidenses foi

[24]Testei cada um dos parâmetros (duração de VOT, oclusão e consoante) para cada nacionalidade nas duas línguas. As tabelas com os 24 testes se encontram nos Apêndices 7 e 8.

de *U*=0; *p*=0,01, as durações absolutas de oclusões das brasileiras apresentaram *U*=0; *p*=0,02 e as durações de oclusões relativas das brasileiras apresentaram *U*=5,5; *p*=0,017.

Com relação às durações totais de consoantes, as durações absolutas das estadunidenses em inglês foram os únicos resultados significativamente diferentes, com *U*=2; *p*=0,02, comparando textos e sentenças-veículo. Já em português, os únicos resultados que não apontaram significância entre textos e sentenças-veículo foram as durações relativas das estadunidenses. Em português, a duração absoluta das estadunidenses foi de *U*=6; *p*=0,38, as durações absolutas totais de consoantes das brasileiras apresentaram *U*=0; *p*=0,02 e as durações relativas totais de consoantes das brasileiras apresentaram *U*=7,5 *p*=0,034.

Com os resultados apresentados anteriormente, é possível afirmar que houve variabilidade nos resultados entre os diferentes parâmetros observados entre os testes, porém não é possível confirmar a hipótese de que existem diferenças entre os testes. Uma pesquisa especificamente desenhada para essa finalidade precisa ser desenhada para responder essa inquietação.

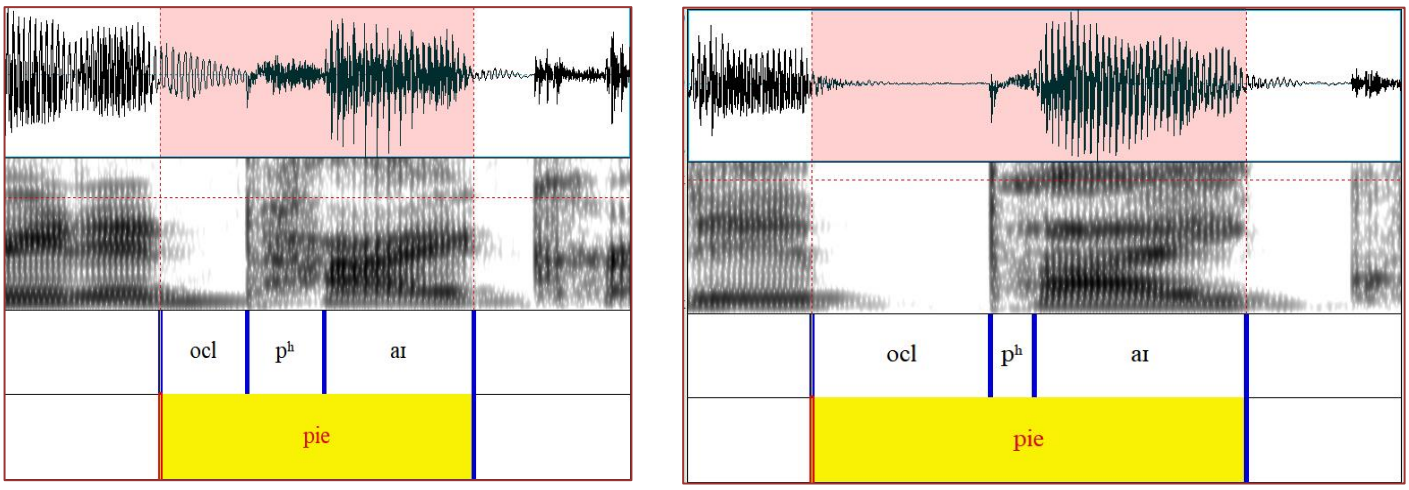
Dedico a próxima seção às análises qualitativas.

6.2. ANÁLISE QUALITATIVA

Para efetuar as análises qualitativas, selecionei duas informantes, cujas coletas apresentaram poucos ruídos e nenhum dado foi descartado por eventuais erros nas leituras ou perda de arquivos. As informantes foram uma americana e uma brasileira, nomeadas, respectivamente, de A3 e B15. A seguir, disponho espectrogramas lado a lado para comparações qualitativas. As primeiras comparações, representadas na Figura 31, são referentes às produções em inglês (primeiramente por meio da tarefa de texto, em seguida por meio da tarefa de leitura de sentenças-veículo)

FIGURA 31 – PRODUÇÕES DAS INFORMANTES EM INGLÊS. À ESQUERDA A3 NO TEXTO, À DIREITA, B15 NO TEXTO

/Ocl 75ms VOT 74ms Cons 149ms/ [25] /Ocl 142ms VOT 35ms Cons177ms/

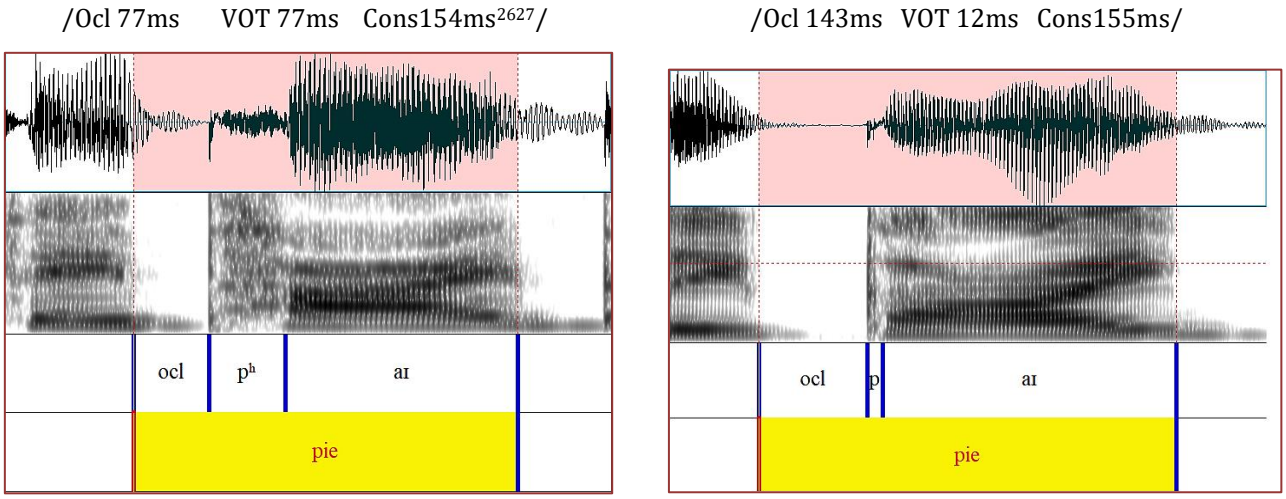


FONTE: O autor (2021).

[25]Disponho, entre colchetes, a duração dos segmentos de duração de Oclusão, VOT e total da consoante.

Na Figura 31, é possível observar a grande diferença no período de oclusão – a duração de oclusão da informante brasileira (em uma proporção de oclusão sendo quase quatro vezes o valor do VOT) é quase o dobro da duração de oclusão da americana (em que o VOT e a oclusão possuem praticamente a mesma duração). O VOT da americana é mais que o dobro do VOT da brasileira e é visível o avantajado setor do espectrograma com ruído (correspondente ao VOT) que a americana possui em comparação com a brasileira. Chamo atenção também para a barra de vozeamento da fala das informantes: O período de oclusão de A3 é menor, porém com uma barra de vozeamento (que esmaece na articulação da plosiva); o resquício de vozeamento da falante brasileira é bem menor, e uma grande parte da barra de vozeamento do período de oclusão é *limpo,* ou seja, sem vozeamento. Na Figura 32, continuo comparando as produções de A3 e B15 em inglês, porém agora via sentenças-veículo.

FIGURA 32 – PRODUÇÕES DAS INFORMANTES EM INGLÊS. À ESQUERDA A3 NAS SENTENÇAS, À DIREITA, B15 NAS SENTENÇAS-VEÍCULO



FONTE: O autor (2021).

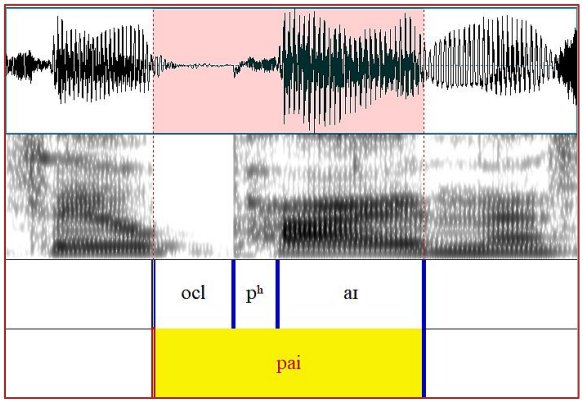
Observando-se a Figura 32, nota-se que alguns padrões da Figura 31 se mantêm: a duração da oclusão da informante brasileira (mais de 10 vezes maior que o VOT continua sendo, praticamente, o dobro da oclusão da americana, que possui uma duração igual ao VOT, e o VOT da americana continua muito maior que o da brasileira – aqui sendo mais que o triplo. A diferença visual de um VOT na casa dos 70ms e um VOT na casa dos 10ms é nítida (precisei inserir o /p/ na imagem da direita manualmente, de tão pequena que é a duração). A barra de vozeamento no período de oclusão de A3 continua mais presente que para B15 e o mesmo fenômeno de esmaecimento notado na Figura 31 ocorre para a brasileira. Nota-se também que a duração total da consoante é 1ms diferente, ou seja, praticamente igual, fato que fez emergir o padrão de proporção entre duração de VOT grande/ oclusão curta, referente ao questionamento se as durações totais de consoantes

[26] As imagens foram reduzidas de tamanho para poder facilitar a comparação com uma observação lado a lado.
[27] Para melhor qualidade das imagens, extraí todos os arquivos de áudio para *mono* em vez de *stereo.*

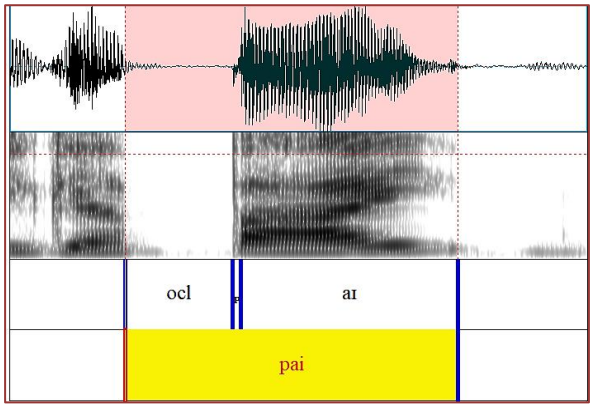
seriam iguais entre as línguas e nacionalidades. Na Figura 33, apresento as diferenças de espectrograma entre A3 e B15 nos textos em português.

FIGURA 33 – PRODUÇÕES DAS INFORMANTES EM PORTUGUÊS. À ESQUERDA A3 NO TEXTO, À DIREITA, B15 NO TEXTO

/Ocl 82ms VOT 45ms Cons 127ms/



/Ocl 146ms VOT 12ms Cons 158ms/

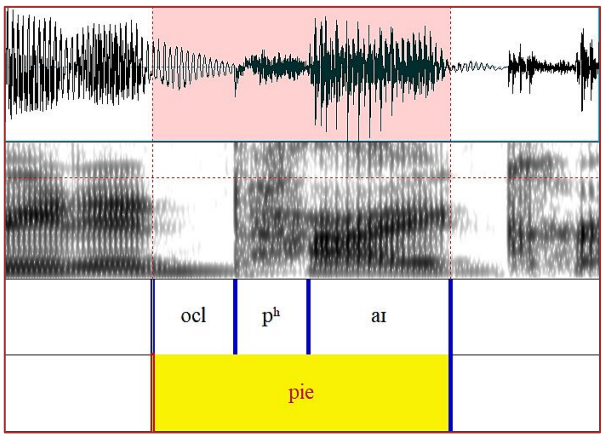


FONTE: O autor (2021).

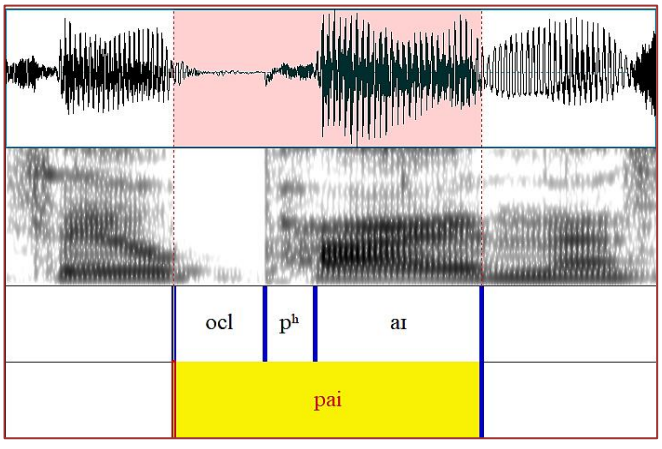
Na Figura 33, ainda é possível encontrar diferenças no tempo de duração de VOT e de oclusão. A duração de VOT de A3 é um pouco maior que o triplo de B15 (45ms e 12ms, respectivamente) e a oclusão da B15, por sua vez, continua bem maior do que a da informante A3 (146ms e 82ms, respectivamente). A questão da barra de vozeamento aqui não diferencia as informantes, uma vez que ambas esmaecem como a brasileira havia feito nos exemplos das Figuras 31 e 32. Nos testes estatísticos apresentados na Seção 6.1, as maiores diferenças significativas realmente haviam ocorrido na língua inglesa, demonstrando que as produções em português se mostram mais semelhantes. Na Figura 34, utilizo as produções em inglês e português da informante A5.

FIGURA 34 – PRODUÇÃO DA INFORMANTE A3 NAS DUAS LÍNGUAS. À ESQUERDA A3 EM INGLÊS E À DIREITA A3 EM PORTUGUÊS

/Ocl 75ms VOT 74ms Cons 149ms/



/Ocl 82ms VOT 45ms Cons 127ms/



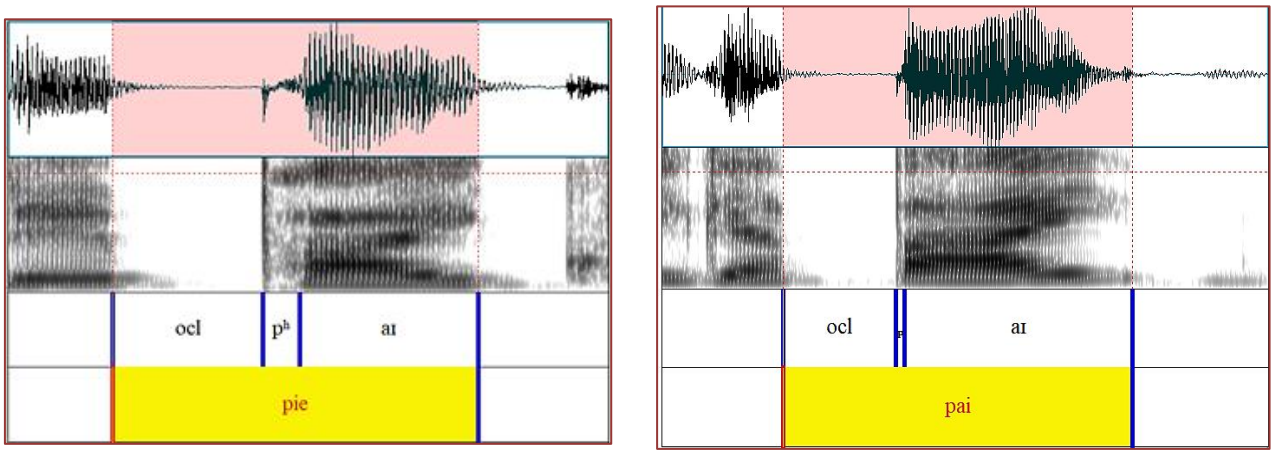
FONTE: O autor (2021).

A Figura 34 compara a mesma informante estadunidense nas duas línguas estudadas: português e inglês. É possível observar a tendência consolidada pela literatura de um VOT maior em inglês (no caso 74ms em inglês e 45ms em português). O setor do VOT em inglês aparece muito mais ruidoso do que em português, quiçá pela duração maior da aspiração. A forma de onda do ruído em inglês também aparenta ter maior amplitude do que em português. Nessa figura, o período de oclusão em inglês pode ser considerado vozeado, enquanto em português há um vozeamento parcial no começo do período de oclusão e um esmaecimento antes da soltura da oclusiva.

Na Figura 35, comparo a informante brasileira nas duas línguas (tarefa de texto).

FIGURA 35 – À ESQUERDA, B15 NO TEXTO EM INGLÊS. À DIREITA, B15 NO TEXTO EM PORTUGUÊS

/Ocl 142ms VOT 35ms Cons 177ms/ /Ocl 146ms VOT 12ms Cons 158ms/



FONTE: O autor (2021).

É interessante observar a grande diferença de VOT da informante brasileira nas duas línguas. A diferença da americana na Figura 34 já era notável – a da brasileira (sendo que o VOT é, praticamente, três vezes menor em português) é ainda mais perceptível. Nota-se também que as oclusões são, praticamente, do mesmo tamanho e com o mesmo vozeamento inicial nas duas imagens, indicando uma possível característica da informante.

As Figuras 31, 32, 33, 34 e 35 apresentaram produções de duas informantes em línguas e instrumentos de pesquisa diferentes. A partir desses dados, apontam-se algumas possíveis contribuições à área de comparação de locutor. Ao observar espectrogramas individualmente, alguns detalhes que fogem à simples média de duração quantitativa se apresentam. A questão do esmaecimento da barra de vozeamento durante o período de oclusão pela informante B15 em várias ocasiões e algumas presenças da barra pela A3 pode indicar a tendência maior de produção desse vozeamento de oclusão por B15 em detrimento de A3. Ao olhar para os casos individuais, espera-se encontrar períodos oclusivos maiores para brasileiras e para a língua portuguesa, além de períodos de VOT maiores para estadunidenses e para a língua inglesa. A Figura 33 evidencia a diferença de VOT de 45ms para 12ms – os VOTs abaixo de 15ms nem sequer apresentam a etiquetagem de /p/ em uma janela temporal de observação de 550-700ms de duração. É possível

também observar as diferenças nas durações de VOT ao olhar para uma mesma informante nas duas línguas.

Para além da questão da consoante /p/, cabe observar que a informante A3 apresentava *creaky voice,* ou voz crepitante, percebida na vogal seguinte à consoante de ataque. Esse fato pode ser observado, principalmente, na Figura 31 à esquerda, por meio de interferências verticais nos espectrogramas que denotam os *cracks* na voz. Verifica-se que a grande maioria dos casos de voz crepitantes observados durante a etiquetagem pertenciam a estadunidenses, porém como não foram feitas estatísticas por não ser o foco da pesquisa, trago isso a título de curiosidade. A língua é complexa e variável, não são alguns poucos segmentos ou subsegmentos de sons que possuem o poder inequívoco de descriminar locutores. Espera-se que cada estudo ajude que a fala seja cada vez mais mapeada e conhecida e que este estudo de durações de oclusões, VOTs e durações totais de consoantes tenha colaborado.

6.3. VERIFICAÇÃO DAS HIPÓTESES

Na Seção 5.3, havia apresentado quatro hipóteses para a pesquisa. Após a apresentação dos dados, é necessário retomar as proposições levantadas e traçar conclusões sobre elas.

(1): O grupo de estadunidenses produzirá oclusões significativamente mais curtas, VOTs significativamente mais longos e consoantes parecidas com o grupo de brasileiras em ambas as línguas.

As durações absolutas/relativas de VOT das estadunidenses em inglês foram de 72,6ms/18,1%, enquanto o das brasileiras foram de 47,6ms/11,5%. Os resultados **foram** estatisticamente significativos. A duração absoluta/relativa do VOT das estadunidenses em português foi de 26ms/6,3%, enquanto o das brasileiras foi de 14,9ms/4,4%. As diferenças **não foram** estatisticamente significativas. As durações absolutas/relativas da oclusão das estadunidenses em inglês foram de 92,2ms/23,3%, enquanto as das brasileiras foi de 115,2ms/28,0%. Ambos os valores **foram** significativamente diferentes entre si. A durações absoluta/relativa da oclusão das estadunidenses em português foram de 106,9ms/26,3%, enquanto o das brasileiras foi de 119,4ms/32,8%. As diferenças **não foram** estatisticamente significativas na duração absoluta, mas foram para a duração relativa. Os valores das durações totais da consoante para as estadunidenses em inglês foram de 164,8ms/41,8%, enquanto os das brasileiras foram de 162,4ms/39,5%. O valor da duração total da consoante para as estadunidenses em português foi de 132,9ms/32,9%, enquanto o das brasileiras foi de 134,3ms/37,2%. Diferenças significativas **foram** observadas somente nas durações relativas totais em português.

A primeira hipótese foi parcialmente confirmada (em casos como a duração absoluta do VOT em inglês e duração absoluta do VOT em inglês, por exemplo). Nota-se que os dados são muito complexos para uma generalização tão grande como a proposta pela primeira hipótese.

(2) O grupo de estadunidenses produzirá, em inglês, oclusões significativamente mais curtas, VOTs significativamente mais longos e a duração total da consoante será parecida na comparação com o português.

As durações absolutas/relativas de VOT das estadunidenses em inglês foram de 72,6ms/18,1%, já em português, os resultados foram de 26ms/6,3% (**significativamente**

diferentes). As durações absolutas/relativas da oclusão das estadunidenses em inglês foram de 92,2ms/23,3%. Em português, foram de 106,9ms/26,3%, com a duração absoluta **significativamente diferente.** Os valores das durações totais da consoante para as estadunidenses em inglês foi de 164,8ms/41,8%, já em português foi de 132,9ms/32,9%. Aqui, somente a duração relativa foi **significativamente diferente.**

A segunda hipótese também foi parcialmente confirmada, com durações de VOT e oclusão diferentes entre as línguas para o grupo de estadunidenses.

(3) O grupo de brasileiras produzirá, em inglês, oclusões significativamente mais curtas, VOTs significativamente mais longos e a duração total da consoante será parecida na comparação com o português.

As durações de VOT em inglês para brasileiras foram de 47,6ms/11,5%, já em português foram de 14,9ms/4,4% (**durações absolutas e relativas foram significativamente diferentes).** No que se refere à oclusão, as durações das brasileiras em inglês foram de 115,2ms/28,0%, enquanto em português foram de 119,4ms/32,8% (**não diferentes, estatisticamente)**. Já nas durações totais da consoante, as brasileiras apresentaram 162,4ms/39,5% em inglês e 134,3ms/37,2% em português (**durações absolutas totais de consoantes significativamente diferentes)**.

A terceira hipótese foi parcialmente confirmada, com as brasileiras produzindo de forma significativamente diferente entre as duas línguas duração de VOT absoluto e duração absoluta total da consoante, mas não o valor de oclusão absoluta entre as duas línguas.

(4) Haverá diferenças significativas entre as tarefas de texto e sentenças-veículo ao observar, em ambas nacionalidades e línguas, durações de oclusão, VOT e duração total da consoante.

10 dos 24 testes estatísticos para hipótese 4 apresentaram resultados significativos. Isso representa 41,67%, ou um pouco menos da metade. Como descrito pela hipótese 4, o esperado era que os testes sempre apresentassem diferenças significativas, uma vez que o ritmo de leitura do texto aparentava ser mais rápido do que o ritmo de leitura das sentenças-veículo (e a chance de as palavras nas sentenças-veículo serem enfatizadas era grande). Todavia, observa-se mais uma vez uma grande variabilidade entre os resultados. Apenas uma, das oito possibilidades de comparação com VOT foi significativa, e era das estadunidenses em sua língua materna. Nas oclusões, por sua vez, cinco resultados foram significativos, mas três não o foram.

É possível concluir que não se pode prever regras sobre os instrumentos a partir deste universo de dados e informantes, uma vez que a variabilidade é alta demais entre os parâmetros analisados e entre valores de duração absolutos e relativos. Sempre que houve variabilidade entre os instrumentos de pesquisa, as durações do texto foram mais curtas que as durações das sentenças. Todavia, foram em apenas 10 das 24 comparações que as diferenças foram significativas. **Hipótese parcialmente confirmada para 10/24 comparações.**

Passo agora para as considerações finais do trabalho.

CAPÍTULO

07 CONSIDERAÇÕES FINAIS

O presente capítulo está dividido de forma sequencial: retomarei alguns pontos desde o Capítulo 2 até o Capítulo 6.

O Capítulo 2 deste trabalho contou com uma descrição dos Sistemas Complexos, seja a denominação mais atual, "Dinâmicos" (SDC), ou a denominação "Adaptativos" (SAC). A complexidade foi a visão teórica de língua de base para o presente trabalho. Selecionei certos temas referentes à complexidade e aos SDCs para tratar de forma mais filosófica e reflexiva no que se refere à visão de língua. Como já mencionado no Capítulo 2, vários dos tópicos levantados por Baranger (2000) possuem intersecções com os pontos abordados por Chan (2001) e Rzevski (2012). Alguns, no entanto, só se encontram nos últimos dois autores. O primeiro tópico trazido por Chan (2001), segundo trazido por Rzevski (2012), trata do controle distribuído. As regras implícitas na fala cumprem essa função. Entretanto, não existe um controle central que define quais produções são as que devem ser obrigatoriamente produzidas. A língua falada na rua pode divergir do que é encontrado nas gramáticas normativas. As produções de VOT observadas, conforme descritas na Seção 6.1 e nos parágrafos anteriores, seguem um fio e saem pela tangente. Talvez, com um *corpus* maior ou observando os informantes ao longo do tempo, como em Schereschewsky (2021), esse aspecto poderia ser mais bem observado. O sétimo tópico trabalhado por Rzevski (2012), terceiro levantado por Chang (2001), trata da coevolução de um SDC. Aqui esbarro, novamente, na impossibilidade de coletar dados das informantes deste trabalho nos meses ou anos depois da gravação original, com o intuito de observar, diacronicamente, as mudanças em todos os parâmetros analisados. Certamente, a cada dia de fala, a cada uso da L1 e da L2, subsistemas podem estar coevoluindo. Apesar de não ter feito uma análise diacrônica, afirmo que a análise sincrônica realizada, esse pequeno recorte no tempo, poderia ser diferente se fosse feito um ano depois ou um ano antes. Neste trabalho, foram observadas algumas variações da mesma informante ao analisar produções nas duas línguas (Seção 6.2). Com relação ao quarto tópico trazido por Baranger (2000), acerca da capacidade de um SDC fazer emergir comportamentos, havia exemplificado o fato de ser impossível olhar a ação do corpo humano se movimentar observando apenas uma perna isoladamente. Isso justifica a investigação das medições de duração da oclusão e da duração total da consoante: como disseram Cristófaro-Silva *et al.* (2019), a consoante oclusiva é composta de seis momentos e não apenas da oclusão e do VOT. Logo, ampliar a área de observação pode trazer *insights* maiores sobre as possíveis diferenças entre informantes em se tratando especificamente da consoante /p/.

O Capítulo 3 buscou descrever, brevemente, o campo de estudo da Fonética e os motivos pelos quais a área é importante. Após a descrição, o campo da forense foi apresentado, com um breve histórico da área e do desenvolvimento dos termos verificação de falante, identificação de falante e, por fim, comparação de locutor. Demonstrei as possíveis contribuições deste trabalho para a área da forense na Seção 6.2,

com considerações após cada apresentação de estatísticas descritivas e inferenciais. Dentre as possíveis contribuições mencionadas, estão a coleta de dados nas duas línguas quando o sujeito da investigação for um falante bilíngue – o VOT e a oclusão sendo fatores que separam grupos de nacionalidades (nesse caso sendo somente entre estadunidenses da costa leste ou do norte dos Estados Unidos vivendo em Curitiba, na região Sul do Brasil e brasileiras nativas e vivendo na mesma região brasileira) servindo como corte/gargalo para um número maior de sujeitos. Há também as diferenças dentro de cada nacionalidade, uma vez que estadunidenses, por exemplo, produziram durações absolutas de VOT e durações absolutas de oclusões, significativamente, diferentes em inglês e português. O desvio-padrão entre as produções em língua materna e língua estrangeira também se mostram importantes fatores de diferenciação de nacionalidade, uma vez que demonstram que a variabilidade em uma língua estrangeira aparenta ser maior. O presente trabalho também pode contribuir para tópicos como os debatidos por Foulkes *et al.* (2019), que apresentam a análise linguística para a determinação de origem (LADO) como perfilamento de locutores em um contexto forense. Analisar as durações de VOT, oclusão e durações totais da consoante de uma informante pode ser útil para o almejado perfilamento da fala de um locutor. As variações de VOT e oclusão nas falas de um locutor, por exemplo, podem contribuir para a definição do perfil sociolinguístico e/ou educacional do indivíduo (uma brasileira com durações muito altas de VOT pode indicar que houve desenvolvimento de um outro subsistema linguístico que pode ter sido adquirido por meio do estudo no Brasil ou de uma imersão linguística em um país anglófono).

O Capítulo 4 revisou consoantes oclusivas, período de oclusão e VOT. A presente pesquisa é mais um pequeno tijolo na enorme construção do conhecimento linguístico, já que adiciona mais valores de base para durações de oclusões, VOTs e durações totais de consoantes para brasileiras e estadunidenses bilíngues, com o diferencial de as estadunidenses viverem no Brasil. Foi mostrado que o VOT da língua inglesa, realmente, é mais longo que o da língua portuguesa para ambas as nacionalidades; durações totais de consoantes foram muito parecidas nas duas línguas entre as nacionalidades; maiores em inglês e menores em português. Além disso, uma importante contribuição do trabalho atual é a descrição de dados de duração de oclusão e o efeito compensatório da consoante em falantes bilíngues brasileiras e estadunidenses nas línguas portuguesa e inglesa. A área de oclusão não possui tantas referências como a de VOT e espero que os resultados apresentados aqui possam ter colaborado para o ramo.

A grande dificuldade da metodologia do trabalho, apresentada no Capítulo 5, foi a pandemia de coronavírus que paralisou as atividades acadêmicas no meu primeiro mês de mestrado. Alguns bancos de dados foram gentilmente cedidos a mim e precisei escolher com qual trabalhar e o que analisar. Após etiquetar todas as ocorrências de todos as oclusivas para duas informantes, o foco precisou ser colocado na consoante /p/ pela maior similaridade da palavra *pie* e pai nos textos e sentenças-veículo em inglês e português. Agradeço demais pelos dados emprestados, uma pequena parte na qual já havia trabalhado no escuto de ditongos há uns anos. Nesses cinco anos de Fonética e Fonologia, ainda não coletei meu banco de dados, mas fico ansioso por essa experiência que ocorrerá durante o Doutorado.

Diversos resultados significativos foram encontrados no Capítulo 6. Na Seção 6.1, descobriu-se que estadunidenses e brasileiras produziram de formas significativamente diferentes durações absolutas/relativas de VOT, durações absolutas/relativas de oclusões e durações relativas totais de consoantes em língua inglesa, além de durações de

oclusões relativas em português. Além disso, verificou-se que estadunidenses produziam durações absolutas/relativas de VOTs, durações absolutas de oclusões e durações relativas totais de consoantes significativamente diferentes entre as duas línguas. As brasileiras, por sua vez, apresentaram durações absolutas/relativas de VOTs, oclusões relativas e durações absolutas totais de consoantes significativamente diferentes entre as línguas. Uma possível continuação de pesquisa que se interliga à Seção 6.2 de análises qualitativas está ligada ao período de oclusão descrito por Abramson e Whalen (2018). Os autores vão além da simples descrição de *Período de Oclusão* e etiquetam os dados por duração de oclusão vozeada e não vozeada. Esse parâmetro da barra de vozeamento durante o período de oclusão pode ser um fator individualizante ao observar informantes como A3 e B15, uma vez que a informante A3 apresentou períodos de oclusão vozeada geralmente mais longos. Na Seção 6.3 do capítulo de resultados, que buscou retornar hipóteses, verifiquei que todas as hipóteses iniciais foram parcialmente comprovadas: dentre alguns dados que demonstraram diferenças significativas, destaco o fato de ambas as nacionalidades produzirem, à sua maneira, VOTs mais longos em inglês do que em português. Isso corrobora o status da literatura que afirma VOTs em inglês mais longos (na presente pesquisa 72,6ms para estadunidenses e 47,6ms para brasileiras). Noto também que nenhuma hipótese foi totalmente refutada. Isso enfatiza a importância de analisar informantes individualmente e qualitativamente para observar em que espaço se encontram as produções de indivíduos específicos, e não apenas se guiar por médias estabelecidas pela literatura, principalmente em um contexto forense. Talvez os resultados fossem diferentes se as hipóteses fossem menos abrangentes e mais específicas. Nota-se que a grande variabilidade dos dados fortalece a adoção da visão de língua como SDC.

Os objetivos e hipóteses foram escritos e reescritos diversas vezes entre a qualificação, defesa e versão final, e poderiam ser melhor elaboradas ainda. Nas versões finais, acabaram ficando generalistas demais (abarcando oclusão, VOT e consoante e não demonstrando completude de ser ou não significativo em todos os parâmetros nenhuma vez). Acredito que a experiência de delimitar objetivos e hipóteses tenha sido válida e certamente vai me habilitar para tecer de forma mais inteligente esta importante parte do trabalho em uma futura experiência de Doutorado.

Outras parcelas do trabalho que poderiam ter sido mais bem aprofundadas e planejo publicar na forma de artigos são: uma investigação mais detalhada em atrito linguístico e a análise da possível ocorrência de atrito em informantes de ambas as nacionalidades; uma discussão sobre as discrepâncias entre valores absolutos e relativos de durações de VOT, oclusões e durações totais de consoantes. Isso pode ser importante para área, uma vez que diversos trabalhos apenas realizam medições absolutas; uma análise mais individualizada e comentada das informantes de forma individual, além de uma análise mais detalhada sobre os gráficos de dispersão; uma leitura mais aprofundada nos trabalhos recentes de SDC que abordam temas como análise de produto, análise de processo.

De forma geral, espero que o presente trabalho possa contribuir para a área da Fonética Acústica mesmo tendo sido desenvolvido em tempos tão complexos e com tantos desafios acadêmicos e não acadêmicos tendo sido colocados. Também desejo que mais pesquisas investigando desenvolvimentos diacrônicos de oclusão e VOT, padrões em brasileiras que moram nos Estados Unidos e estudos de percepção, por exemplo, sejam desenvolvidos como ramificações da atual pesquisa.

REFERÊNCIAS

[1] ABRAMSON, A. S.; WHALEN, D.H. Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. **Journal of Phonetics**, 63, p. 75-86, 2018. Disponível em:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5665574/>. Acesso em: 20 out. 2020.

[2] ALBUQUERQUE, J. **Caminhos dinâmicos em Inteligibilidade e Compreensibilidade de Línguas Adicionais**: um estudo longitudinal com dados de fala de haitianos aprendizes de Português Brasileiro. Tese (Doutorado em Linguística) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019

[3] ALVES, M.; SEARA, I. Produção de plosivas surdas em inglês e português por falantes brasileiros de inglês como língua estrangeira. **Anais do CELSUL**, 2008.

[4] BALCI, S.; JENTSCHKE, S.; GALLUCCI, M.; SELKER, R.; DROPMANN, D. LOVE, J. **Jamovi** – Stats Open Now. Disponível em: <https://www.jamovi.org/> . Acesso em 23 abr. 2021.

[5] BARANGER, M. **Chaos, Complexity and Entropy: A physics talk for non-physicists.** New England Complex Systems Institute. Abril 2000. Disponível em: http://www.necsi.edu/projects/baranger/cce.html. Acesso em: 27 jul. 2020.

[6] BARBOSA, P.A.; MADUREIRA, S. **Manual de fonética acústica experimental - Aplicações a dados do português**. São Paulo: Editora Cortez, 2015.

[7] BEREJUK, D.; GOMES, M. No prelo.

[8] BERNARDO-SILVA, O. **Análise Acústica da produção de fricativas alveolares em coda na língua inglesa por falantes bilíngues brasileiras e estadunidenses sob uma perspectiva de Sistemas Adaptativos Complexos**. TCC (Trabalho de Conclusão de Curso) – Universidade Tecnológica Federal do Paraná, Curitiba, 2019.

[9] BERNARDO-SILVA, O.; GOMES, M. Análise acústica do ditongo /eɪ/ em inglês e em português por falantes bilíngues. **Language and Law**, v. 6, n.1, 2019.

[10] BOERSMA, P.; WEENINK, D. **Praat**: Doing phonetics by computer**.** Software computacional. Disponível em: <https://www.fon.hum.uva.nl/praat/>. Acesso em: 02 mar. 2021.

[11] BRITISH COUNCIL, 2014. **Learning English in Brazil**. Disponível em: <https://www.britishcouncil.org.br/sites/default/files/learning_english_in_brazil.pdf >. Acesso em 09 abr. 2021.

[12] CASTRO, T.; ALBUQUERQUE, J.; GOMES, M. A produção de vogais médias do Português como Língua Adicional: uma análise de dados preliminares de um aprendiz chileno via sistemas dinâmicos complexos. **ReVEL**. v. 18, n. 35, 2020. Disponível em: < /www.revel.inf.br/ >. Acesso em: 21 maio 2021.

[13] CELESTE, L. C.; TEIXEIRA, E. Efeito do sexo e idade na produção do VOT. **Revista de Letras da Universidade Católica de Brasília**, Ano II, v. 2, n. 1, jul. 2009.

[14] CHAN, S. Complex Adaptive Systems. **Research Seminar in Engineering Systems***.* 2001.

[15] CHEUNG, W.; WEE, L. Viability of VOT as a parameter for speaker identification: evidence from Hong Kong. In: **18h International Congress of Linguistics**. p.3921-3934. 2008.

[16] CHO, T.; LADEFOGED, P. **Variation and universals in VOT**: evidence from 18 languages. Phonetics Lab, Department of Linguistics, UCLA, Los Angeles, CA 90095, U.S.A. 1999.

[17] COHEN, G. **The VOT dimension**: A bi-directional experiment with English and Brazilian-Portuguese stops. Dissertação (Mestrado em Linguística) – Universidade Federal de Santa Catarina. Florianópolis, 2004.

[18] CRISTÓFARO-SILVA, T. **Fonética e Fonologia do Português: Roteiro de Estudos e Guia de Exercícios**. Editora Contexto. 7 ed. São Paulo, 2003.

[19] CRISTÓFARO-SILVA, T; SEARA, I; SILVA, A. H. P; RAUBER, A. S.; CANTONI, M. **Fonética Acústica: os sons do português brasileiro**. São Paulo: Editora Contexto, 2019.

[20] CRYSTAL, T.; HOUSE, A. **The duration of American-English stop consonants: an overview.** Communication Research Division, Institutute for Defense Analyses, Princeton, Estados Unidos da América. 1987.

[21] DE BOT, K. Complexity Theory and Dynamic Systems Theory: Same or Different? In **Studies in Second Language Learning and Teaching**, v. 48, 2017.

[22] DE BOT, K.; LOWIE, W.; VERSPOOR, M. A Dynamic Systems Theory approach to second language acquisition. **Bilingualism**: Language & Cognition, p. 7-21, 2007.

[23] DIAS, M. D. S.; GODINHO, C.; P.; PACHECO, V. A produção das consoantes oclusivas do inglês por falantes nativos e brasileiros: a relação entre duração e soltura. **Revista do GELNE**, Natal/RN, v. 18, n. 1, p. 93-115, 2016.

[24] FLEGE, J.; BIRDSONG, D.; BIALYSTOK, E. MACK, M. SUNG, H. TSUKADA, K. Degree of foreign accent in english sentences produced by korean children and adults**. Journal of Phonetics**, n. 34. p.153-175, 2006.

[25] FLEGE, J.; BOHN, O. **The revised Speech Learning Model (SLM-r).** 2021.

[26] FLEGE, J.; PISKE, T.; MACKAY, I.; MEADOR, D. The production of English vowels by fluent early and late Italian-English bilinguals. **Phonetica**, v. 59, p. 49-71, 2002.

[27] FLEGE, J.; SCHIRRU, C.; MACKAY, I. Interaction between the native and second language phonetic subsystems. **Speech Communication**, n. 40, p. 467-491, 2003.

[28] FOULKES, P.; FRENCH, P.; WILSON, K. LADO as a forensic speaker profiling. In: **Language Analysis for the Determination of Origin**. Springer International Publishing, p.91-116, 2019.

[29] FOULKES, P.; SCOBBIE, J.; WATT, D. Sociophonetics.. Em **The Handbook of Phonetics Science**. Blackwell Publishers, MA, Estados Unidos da América. 2010.

[30] GOMES, M.; CARNEIRO, D.; DRESCH, A. Análise perceptiva e acústica em fonética forense: uma pesquisa em disfarce de voz. **Domínios da Lingu@.** p. 559-589, 2016.

[31] GOMES, M. Normalização de frequência de formantes e análise com dados brutos na visão de língua como um Sistema Adaptativo Complexo. **Gradus-Revista Brasileira de Fonologia de Laboratório**, v. 3, n. 1, p. 81-102, 2018.

[32] GONÇALVES, C. **Taxa de elocução e de articulação em corpus forense do português brasileiro**. Tese (Doutorado em Linguística) – Pontifícia Universidade Católica do Rio Grande do Sul. Porto Alegre, 2013.

[33] HOLLAND, J. Studying Complex Adaptive Systems. **Journal of Systems Science and Complexity**. 2006.

[34] HOLLIEN, H. Forensic Voice Identification. **Academic Press**, Estados Unidos. 2002.

[35] JESSEN, M.; O. GFROERER, S. Influence of vocal effort on average and variability of fundamental frequency**. International Journal of Speech Language and the law.** p.174-213. 2015.

[36] KAMPFF, F.R. **Desenvolvimento de padrões de Voice Onset Time (VOT) em plosivas iniciais do inglês (l2) por aprendizes do sul do brasil expostos ou não a treinamento perceptual**: uma verificação longitudinal em quatro etapas. TCC (Trabalho de Conclusão de Curso) – Universidade Federal do Rio Grande do Sul. Porto Alegre, 2018.

[37] KENT, R.D.; READ, C. **The Acoustic Analysis of Speech**. Connecticut, USA: Thomson, 2011.

[38] KUPSKE, F.F. **Imigração, atrito e complexidade: a produção das oclusivas surdas iniciais do inglês e do português por sul-brasileiros residentes em Londres**. Tese (Doutorado em Linguística) – Universidade Federal do Rio Grande do Sul. Porto Alegre, 2016.

[39] LADEFOGED, P.; JOHNSON, K. **A Course in Phonetics**. Wadsworth, Cengage Learning. University of California. Los Angeles, 2011.

[40] LARSEN-FREEMAN, D. Complexity Theory: A New Way to Think. **RBLA**, Belo Horizonte, v.13, n. 2, p. 269-373, 2013

[41] LIMA JR, R. Análise dinâmica da aquisição do das oclusivas do inglês por alunos brasileiros em diferentes idades. **Revista do GELNE**, v. 17, n. ½, p. 55-75, 2015.

[42] LISKER, L.; ABRAMSON A.S. A Cross-Language Study of voicing in initial stops: Acoustical measurements. **Word**, 2003, p.384-422, 1964.

[43] LOFREDO-BONATTO, M.; SILVA, M. Comparação de sons plosivos em crianças monolíngues e bilíngues por meio do parâmetro acústico voice onset time: relato de casos. **Revista CEFAC**, v. 20, p. 680-687, 2018.

[44] LOPES, R.; BAUMGARTNER, C. Inglês como Língua Franca: explicações e implicações. **The especialist.**, v. 40, n. 2, 2019.

[45] LOUSADA, M. **Estudo da Produção de Oclusivas do Português Europeu.** Dissertação (Mestrado em Linguística) – Universidade de Aveiro, 2006.

[46] LOWIE, W.; VERSPOOR, M. H. Individual differences and the ergodicity problems. **Language Learning**, v. 69, n. S1, p.184-206, 2019.

[47] MARTINS, L. S. Bilinguismo (Hunrückisch-português brasileiro) e transferências fonético-fonológicas. In: **Aquisição fonético-fonológica de língua estrangeira**. Alves, U. K. Editora Pontes, 2016.

[48] MELO, R. MOTA; H. MEZZOMO, C.; BRASIL, B.; LOVATTO, L.; ARZENO, L. Acoustic characterization of the voicing of stop phones in Brazilian Portuguese. **Revista CEFAC**, p. 487-499, 2014.

[49] NOLAN, F. **Speaker identification evidence**: its forms, limitations, and roles**.** Reino Unido: University of Cambridge, 2001.

[50] NOLAN, F. Speaker Recognition and forensic phonetics. Em **The Handbook of Phonetics Science**. Blackwell Publishers, MA, Estados Unidos da América. 1999.

[51] PAIVA, V,; CORRÊA, Y. Sistemas Adaptativos Complexos: uma entrevista com Vera Lúcia Menezes de Oliveira e Paiva. **ReVEL**, v. 14, n. 27, 2016. Disponível em: <www.revel.inf.br>: Acesso em: 27 de julho de 2020.

[52] REIS, M.S; KLUGE, D.C.; KOERICH, R.D. Efeitos de treinamento perceptual na produção das plosivas não-vozeadas inglesas por aprendizes brasileiros de inglês. **Anais do CELSUL**, 2008.

[53] RZEVSKI, G. **Self-organization versus control in complex social systems**. The Open University, United Kingdom, 2012.

[54] SCHERESCHEWSKY, L. **Desenvolvimento de *Voice Onset Time* em sistemas multilíngues (Português – L1, Inglês – L2 e Francês L3)**: Discussões dinâmicas a partir de diferentes metodologias de análise de processo. Dissertação (Mestrado em Linguística) – Universidade Federal do Rio Grande do Sul. Porto alegre. 2021.

[55] SCHERESCHEWSKY, L. C.; ALVES, U. K.; KUPSKE. F. F. Atrito de língua materna: os efeitos do inglês (L2) nos padrões de VOT do português brasileiro em um ambiente de L1 dominante. **Revista Digital do Programa de Pós-Graduação em Letras da PUCRS Porto Alegre**, v. 10, n. 2, p. 700-716, jul./dez. 2017.

[56] SCHWARTZHAUPT, B.M. **Testing intelligibility in English**: The effects of positive VOT and contextual information in a sentence transcription task. Dissertação de Mestrado. Universidade Federal do Rio Grande do Sul. Porto Alegre, 2015.

[57] SKARNITZL, R; VANKOVÁ, J**.** Fundamental frequency statistics for male speakers of Common Czech. **AUC Philologica**, p.7-27, 2017.

[58] STATISTA. **The most spoken languages worldwide in 2021.** Disponível em: < https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/> . Acesso em: 09 abr. 2021.

[59] STATHOPOULOS, E; WEISMER, G. Closure of stop consonants. **Journal of phonetics**, v.11, p.395-400, 1983.

[60] VERSPOOR, M; LOWIE, W; Van DIJK, M. Variability in second language development from a dynamic systems perspective. **Modern Language Journal**, 92, p. 214–231, 2008.

APÊNDICES

APÊNDICE 1 – TEXTOS EM PORTUGUÊS E INGLÊS (TRABALHO EXPLORATÓRIO E PESQUISA PRINCIPAL)

Inglês

1. *A classic apple PIE takes a shortcut with easy unroll-fill refrigerated pie crust. Try this sumptuous Bramley apple PIE - simple to make and great to prepare in advance for a dinner party. This was my grandmother's apple PIE recipe. I have never seen another one quite like it. It will always be my favorite and has won me several first-place prizes*.
2. Installments are recurring income tax PAYments that people have to PAY on certain dates. These are to cover tax that you would normally have to PAY in a lump sum on April 30 of the following year. Installments are not paid in advance; they are paid during the calendar year in which you are earning the taxable income. It is always very good to pay your debts!!
3. HI, would you like to BUY a car with confidence? Find out which finance options are available to you so you can BUY it as soon as you like and pay it over a set period of time. Get all the advice you need to BUY your car and read reviews on the latest cars to be driven by our motoring experts.
4. Tampa BAY is a large natural harbor and estuary connected to the Gulf of Mexico on the west central coast of Florida, comprising Hillsboro BAY, McKay BAY, Old Tampa BAY, Middle Tampa BAY, and Lower Tampa BAY. The surrounding area is home to about four million residents, making it a heavily used commercial and recreational waterway but putting much stress on the bays' ecosystem.
5. You LIE down after a long day of feeling lousy, only to find that you can't stop coughing. Just when you need sleep the most, your cough symptoms get worse and keep you up all night. Why is it that you cough when you LIE down? Cough with lying down is often associated with some degree of congestive heart failure. So, tell your doctor if you cough when you LIE down.
6. Lay out a newsletter in publishing layout view. You can lay out a complex publication in publishing layout view and create elements that appear on all pages by putting them on a master page. After you save the document as a template, you can create a new version of the document that is based on the template, and then revise its content. The template keeps the layout the same.

Português

1. Como você já deve saber, no Brasil o Dia dos Pais é comemorado no segundo domingo de agosto. Esse é um dia muito especial. Você que tem um pai ou alguém que representa o seu pai na sua vida, não deixe de abraçá-lo e dizer o quanto ele é importante para você. Ter seu pai é um privilégio!
2. Você sabe o que significa PEI? PEI é a sigla de Programa de Enriquecimento Instrumental. Para facilitar a explicação, são instrumentos desenvolvidos por psicólogos e educadores para ajudar no desenvolvimento dos alunos e incentivar

seu potencial de aprendizagem. Ou seja, o PEI está presente no ambiente acadêmico também.

3. Um dia desses, me peguei pensando na mania que algumas pessoas têm de acrescentar ao seu vocabulário palavrinhas em inglês. Não consegui chegar a nenhuma conclusão. Talvez seja para demonstrar conhecimento ou apenas por questões estéticas. Este pensamento me ocorreu quando estava no meio de uma conversa com minhas colegas e uma delas foi embora disse em voz alta “bai bai meninas”. Oras, não seria muito mais adequado dizer apenas “tchau” ao invés de “bai”? Espera aí!! “Tchau” é português?
4. A editora Bei está completando 19 anos de idade. Essa editora é conhecida por seu variado catálogo que vai de gastronomia à linguística. Ela também é responsável por grandes publicações acadêmicas e livros de grande importância cultural. O nome Bei vem do sobrenome da família que a fundou. João Carlos Bei foi o fundador dessa editora.
5. Você sabia que tem direito às informações públicas quando bem quiser ou precisar? Quem permite esse direito é a Lei nº 12.527/2011 . Essa lei é conhecida como LAI, que significa Lei de Acesso à Informação. Não podemos esquecer que temos muitos benefícios ao nosso favor, devemos sem dúvida usufruir da LAI para combater os problemas sociais. Quando precisar, não se esqueça de recorrer à LAI.

APÊNDICE 2 – SENTENÇAS EM PORTUGUÊS E INGLÊS (ESTUDO EXPLORATÓRIO E PESQUISA PRINCIPAL). AS PALAVRAS APARECEM EM ORDEM ALEATÓRIA EM CADA UMA DAS TRÊS LEITURAS

Say ***days*** badly.
Say ***lucky*** badly.
Say ***kite*** badly.
Say ***bays*** badly.
Say ***low*** badly.
Say ***guys*** badly.
Say ***lies*** badly.
Say ***gays*** badly.
Say ***ties*** badly.
Say ***tomb*** badly.
Say ***Kay*** badly.
Say ***lays*** badly.
Say ***dies*** badly.
Say ***comb*** badly.
Say ***bobby*** badly.
Say ***tie*** badly.
Say ***sigh*** badly.
Say ***pies*** badly.
Say ***rainy*** badly.

Say ***pay*** badly.
Say ***guy*** badly.
Say ***Tay*** badly.
Say ***buy*** badly.
Say ***pie*** badly.
Say ***day*** badly.
Say ***buggy*** badly.
Say ***guy*** badly.
Say ***those*** badly.
Say ***pays*** badly.
Say ***lay*** badly.
Say ***lie*** badly.
Say ***poppy*** badly.
Say ***sighs*** badly.
Say ***bays*** badly.
Say ***bay*** badly.
Say ***die*** badly.
Say ***false*** badly.
Say ***rock*** badly.

Say ***comb*** badly.
Say ***bay*** badly.
Say ***poppy*** badly.
Say ***false*** badly.
Say ***tie*** badly.
Say ***kite*** badly.
Say ***lucky*** badly.
Say ***lies*** badly.
Say ***those*** badly.
Say ***gay*** badly.
Say ***buggy*** badly.
Say ***lays*** badly.
Say ***dies*** badly.
Say ***low*** badly.
Say ***bobby*** badly.
Say ***guys*** badly.
Say ***sigh*** badly.
Say ***pies*** badly.

Say ***pay*** badly.
Say ***guy*** badly.
Say ***Tay*** badly.
Say ***buy*** badly.
Say ***pie*** badly.
Say ***day*** badly.
Say ***gays*** badly.
Say ***guy*** badly.
Say ***Kay*** badly.
Say ***boot*** badly.
Say ***rainy*** badly.
Say ***tie*** badly.
Say ***bays*** badly.
Say ***sighs*** badly.
Say ***rock*** badly.
Say ***tomb*** badly.
Say ***die*** badly.
Say ***days*** badly.

Say ***Kay*** badly.
Say ***sighs*** badly.
Say ***die*** badly.
Say ***Tay*** badly.
Say ***low*** badly.
Say ***rock*** badly.
Say ***boot*** badly.
Say ***comb*** badly.
Say ***pies*** badly.
Say ***gay*** badly.
Say ***buggy*** badly.
Say ***rainy*** badly.
Say ***tomb*** badly.
Say ***kite*** badly.
Say ***lucky*** badly.
Say ***guys*** badly.
Say ***sigh*** badly.

Say ***pay*** badly.
Say ***poppy*** badly.
Say ***tie*** badly.
Say ***buy*** badly.
Say ***pie*** badly.
Say ***day*** badly.
Say ***gays*** badly.
Say ***those*** badly.
Say ***bay*** badly.
Say ***pays*** badly.
Say ***lay*** badly.
Say ***lie*** badly.
Say ***bays*** badly.
Say ***guy*** badly.
Say ***lies*** badly.
Say ***ties*** badly.
Say ***guy*** badly.

Diga **sei** baixo.
Diga **bem** baixo.
Diga **deito** baixo.
Diga **daime** baixo.
Diga **cais** baixo.
Diga **gay** baixo.
Diga **taipa** baixo.
Diga **lei** baixo.
Diga **touca** baixo.
Diga **lema** baixo.
Diga **pais** baixo.
Diga **baita** baixo.
Diga **sais** baixo.
Diga **roupa** baixo.
Diga **sempre** baixo.
Diga **leque** baixo.
Diga **jipe** baixo.
Diga **pai** baixo.
Diga **paica** baixo.
Diga **sou** baixo.
Diga **dei** baixo.

Diga **raiva** baixo.
Diga **peito** baixo.
Diga **lerdo** baixo.
Diga **beiço** baixo.
Diga **sai** baixo.
Diga **paz** baixo.
Diga **teima** baixo.
Diga **queima** baixo.
Diga **tais** baixo.
Diga **seis** baixo.
Diga **leis** baixo.
Diga **reis** baixo.
Diga **cai** baixo.
Diga **deis** baixo.
Diga **sabe** baixo.
Diga **gaita** baixo.
Diga **gays** baixo.
Diga **regue** baixo.
Diga **laico** baixo.
Diga **leito** baixo.
Diga **caibo** baixo.

Diga **regue** baixo.
Diga **laico** baixo.
Diga **dei** baixo.
Diga **gays** baixo.
Diga **cais** baixo.
Diga **gay** baixo.
Diga **taipa** baixo.
Diga **leito** baixo.
Diga **touca** baixo.
Diga **cai** baixo.
Diga **pais** baixo.
Diga **bem** baixo.
Diga **sais** baixo.
Diga **deis** baixo.
Diga **jipe** baixo.
Diga **leque** baixo.
Diga **sei** baixo.
Diga **peito** baixo.
Diga **sou** baixo.
Diga **paica** baixo.
Diga **deito** baixo.

Diga **sempre** baixo.
Diga **teima** baixo.
Diga **lerdo** baixo.
Diga **paz** baixo.
Diga **sai** baixo.
Diga **lei** baixo.
Diga **pai** baixo.
Diga **queima** baixo.
Diga **tais** baixo.
Diga **lema** baixo.
Diga **leis** baixo.
Diga **reis** baixo.
Diga **baita** baixo.
Diga **beiço** baixo.
Diga **sabe** baixo.
Diga **gaita** baixo.
Diga **daime** baixo.
Diga **raiva** baixo.
Diga **roupa** baixo.
Diga **caibo** baixo.
Diga **seis** baixo.

Diga **beiço** baixo.
Diga **sabe** baixo.
Diga **regue** baixo.
Diga **paz** baixo.
Diga **queima** baixo.
Diga **gay** baixo.
Diga **raiva** baixo.
Diga **seis** baixo.
Diga **gaita** baixo.
Diga **cai** baixo.
Diga **pais** baixo.
Diga **laico** baixo.
Diga **leis** baixo.
Diga **deis** baixo.
Diga **gays** baixo.
Diga **leque** baixo.
Diga **tais** baixo.
Diga **daime** baixo.
Diga **roupa** baixo.
Diga **leito** baixo.
Diga **caibo** baixo.

Diga **sempre** baixo.
Diga **teima** baixo.
Diga **lerdo** baixo.
Diga **cais** baixo.
Diga **sai** baixo.
Diga **lei** baixo.
Diga **pai** baixo.
Diga **sais** baixo.
Diga **sei** baixo.
Diga **lema** baixo.
Diga **bem** baixo.
Diga **deito** baixo.
Diga **baita** baixo.
Diga **taipa** baixo.
Diga **peito** baixo.
Diga **jipe** baixo.
Diga **dei** baixo.
Diga **touca** baixo.
Diga **sou** baixo.
Diga **paica** baixo.
Diga **reis** baixo.

APÊNDICE 3 – MÉDIAS DE TODOS OS SEGMENTOS ETIQUETADOS DE A4 NO TEXTO EM INGLÊS (EXPERIMENTO EXPLORATÓRIO)

Segmento	Ocl	VOT	Cons	Ocl R	VOT R	Cons R
/p/ + /a/	53	45	98	12	11	23
/p/ + /i:/ (x2)	105	61	166	24	14	37
/p/+ /ə/ (x3)	40	36	76	7	6	12
/p/ + /u/ (x2)	100	57	157	21	12	32
/t/ + /æ/ (x7)	67	73	140	15	17	33
/t/ + /e/ (x3)	68	57	125	11	9	20
/t/ + /ɛ/	82	61	143	29	22	51
/k/ + /ə/ (x9)	75	70	146	11	12	23
/k/ + /æ/ (x2)	48	64	111	16	20	36
/k/ + /a/ (x3)	69	100	169	16	23	39
/k/ + /ɔ/ (x5)	109	90	199	21	19	40
/k/ + /i:/ (x2)	70	79	149	15	16	31
/k/ + /uai/	62	57	119	23	21	45
/k/ + /ou/	64	86	150	16	22	38
/t/ + /aɪ/	53	100	153	17	32	49
/p/ + /aɪ/ (X4)	75	65	140	20	17	37
/p/ + /ei/ (X9)	84	52	135	23	14	37
Média /p/	76	53	129	18	12	30
Média /t/	67	73	140	18	20	38
Média /k/	71	78	149	17	19	36

APÊNDICE 4 – MÉDIAS DE TODOS OS SEGMENTOS ETIQUETADOS DE A4 NO TEXTO EM INGLÊS (EXPERIMENTO EXPLORATÓRIO)

Segmento	Ocl	VOT	Cons	Ocl R	VOT R	Cons R
/p/ + /a/	78	37	115	18	8	26
/p/ + /i:/ (x2)	67	70	137	21	23	44
/p/+ /ə/ (x3)	64	15	80	10	2	13
/p/ + /u/ (x2)	107	48	155	23	10	33
/t/ + /æ/ (x7)	61	51	112	15	13	28
/t/ + /e/ (x3)	72	45	116	11	7	18
/t/ + /ɛ/	78	62	140	31	25	56
/k/ + /ə/ (x8)	44	46	89	8	9	17
/k/ + /æ/ (x2)	38	83	121	10	28	38
/k/ + /a/ (x3)	66	65	131	19	18	37
/k/ + /ɔ/ (x5)	85	55	139	25	16	41
/k/ + /i:/ (x2)	79	42	121	22	13	35
/k/ + /uai/	28	75	103	11	30	41
/k/ + /ou/	78	65	143	21	17	38
/t/ + /aɪ/	66	82	148	18	22	40
/p/ + /aɪ/ (X4)	89	45	133	25	13	38
/p/ + /ei/ (X8)	103	25	128	24	6	30
Média /p/	85	40	124	20	10	31
Média /t/	69	60	129	19	17	36
Média /k/	60	61	121	16	19	35

APÊNDICE 5 – MÉDIAS INDIVIDUAIS DE DURAÇÕES TOTAIS DAS CONSOANTES EM INGLÊS

Informante	Consoante Total	Consoante Relativa
Estadunidenses	164,8	41,8
Brasileiras	162,8	39,5
A3	157,0	41,0
A4	177,4	43,9
A5	154,9	43,3
A6	172,4	39,9
A7	160,0	42,7
A8	157,4	45,2
A9	174,3	36,7
B7	150,9	40,6
B12	153,3	36,8
B14	156,6	40,2
B15	156,1	36,9
B17	196,7	45,1
B18	190,6	40,7
B20	135,7	36,3

APÊNDICE 6 – MÉDIAS INDIVIDUAIS DE DURAÇÕES TOTAIS DAS CONSOANTES EM PORTUGUÊS

Informante	Consoante Total(ms)	Consoante Relativa(ms)
Estadunidenses	132,9	32,9
Brasileiras	134,3	37,2
A3	118,8	29,9
A4	130,0	36,4
A5	116,7	34,6
A6	123,0	28,3
A7	135,0	31,9
A8	107,2	35,8
A9	200,8	33,2
B7	150,2	39,3
B12	121,8	33,0
B14	124,3	36,2
B15	130,0	36,1
B17	161,0	38,9
B18	137,8	41,8
B20	115,5	35,9

APÊNDICE 7 – TESTES ESTATÍSTICOS PARA AS TRÊS PRIMEIRAS HIPÓTESES. RESULTADOS SIGNIFICATIVOS EM NEGRITO

Testes Independentes (Mann- Whitney)	Valor do Teste	Valor da significância
VOT abs ing ame vs br	***U*=8**	***p*=0,041**
VOT relativo ing ame vs br	***U*=1**	***p*=0,003**
VOT abs PB ame vs br	*U*=17	*p*=0,383
VOT relativo PB ame vs br	*U*=12	*p* =0,128
Ocl abs ing ame vs br	***U*=0**	***p* <0,001**
Ocl relativo ing ame vs br	***U*=6**	***p* =0,017**
Ocl abs PB ame vs br	*U*=15	*p* =0,259
Ocl relativo PB ame vs br	***U*=3**	***p* =0,004**
Cons abs ing ame vs br	*U*=16	*p* =0,318
Cons relativo ing ame vs br	*U*=14	*p* =0,209
Cons abs PB ame vs br	*U*=18,5	*p* =0,482
Cons relativo PB ame vs br	***U*=7**	***p* =0,026**
Testes Pareados (*Wilcoxon*)	Valor do Teste	Valor da significância
VOT abs ame ing vs port	***U*=0**	***p* =0,016**
VOT abs br ing vs port	***U*=2**	***p* =0,047**
VOT relat ame ing vs port	***U*=0**	***p* =0,016**
VOT relat br ing vs port	***U*=0**	***p* =0,016**
Ocl abs ame ing vs port	***U*=28**	***p* =0,016**
Ocl abs br ing vs port	*U*=17	*p* =0,688
Ocl relat ame ing vs port	*U*=19	*p* =0,469
Ocl relat br ing vs port	***U*=28**	***p* =0,022**
Cons abs ame ing vs port	*U*=2	*p* =0,051
Cons abs br ing vs port	***U*=0**	***p* =0,016**
Cons relat ame ing vs port	***U*=0**	***p* =0,016**
Cons relat br ing vs port	*U*=3	*p* =0,078

APÊNDICE 8 – TESTES ESTATÍSTICOS PARA A QUARTA HIPÓTESE. RESULTADOS SIGNIFICATIVOS EM NEGRITO

Parâmetro Analisado	Valor do Teste	Valor da significância
Ame VOT abs ing	***U*=5.5**	***p* =0,018**
Ame VOT relativo ing	*U*=17	*p*=0,369
Bra VOT abs ing	*U*=24	*p* =1
Bra VOT relativo ing	*U*=19,5	*p* =0,561
Ame VOT abs PB	*U*=20	*p* =0,943
Ame VOT relativo PB	*U*=15	*p* =0,427
Bra VOT abs PB	*U*=20,5	*p* =0,654
Bra VOT relativo PB	*U*=14,5	*p* =0,22
Ame Oclusão abs ing	***U*=0**	***p* =0,02**
Ame Oclusão relativa ing	*U*=10,5	*p* =0,082
Bra Oclusão abs ing	***U*=7**	***p* =0,029**
Bra Oclusão relativa ing	*U*=24	*p* =1
Ame Oclusão abs PB	***U*=0**	***p* =0,01**
Ame Oclusão relativa PB	*U*=7	*p* =0,053
Bra Oclusão abs PB	***U*=0**	***p* =0,02**
Bra Oclusão relativa PB	***U*=5.5**	***p* =0,017**
Ame Cons abs ing	***U*=2**	***p* =0,02**
Ame Cons relativa ing	*U*=12	*p* =0,123
Bra Cons abs ing	*U*=9	*p* =0,053
Bra Cons relativa ing	*U*=23	*p* =0,898
Ame Cons abs PB	***U*=6**	***p* =0,038**
Ame Cons relativa PB	*U*=7,5	*p* =0,062
Bra Cons abs PB	***U*=0**	***p* =0,02**
Bra Cons PB	***U*=7,5**	***p* =0,034**

LISTA DE ABREVIATURAS OU SIGLAS

Ame	– Em resultados e tabelas, grupo de estadunidenses
Con	– Nos resultados e tabelas: Duração do período de consoante
Br	– Em resultados e tabelas: grupo de brasileiras
DP	– Desvio-padrão
EUA	– Estados Unidos da América
L1	– Primeira língua/Língua Materna
L2	– Segunda Língua
L3	– Terceira Língua
Ing	– Nos resultados e tabelas: Inglês (nativo estadunidense da costa leste/nordeste ou produzido por brasileiras)
Ocl	– Nos resultados e tabelas: Duração do período de oclusão
PB	– Português Brasileiro (no caso de Curitiba ou região)
RMC	– Região Metropolitana de Curitiba
SAC	– Sistemas Adaptativos Complexos
SDC	– Sistemas Dinâmicos Complexos
SLM	– *Speech Learning Model*
SLM-r	– *Speech Learning Model Revised*
USA	– United States of America
VOT	– *Voice Onset Time*
TSDC	– Teoria dos Sistemas Dinâmicos Complexos

www.poisson.com.br
contato@poisson.com.br

@editorapoisson

https://www.facebook.com/editorapoisson